Strengthening Forensic Science in the United States: A Path Forward

Committee on Identifying the Needs of the Forensic Sciences Community,   National Research Council

ISBN: 0-309-13131-6, 352 pages, 6 x 9, (2009)

This PDF is available from the National Academies Press at:
http://www.nap.edu/catalog/12589.html

Visit the National Academies Press online, the authoritative source for all books from the National Academy of Sciences, the National Academy of Engineering, the Institute of Medicine, and the National Research Council:

- Download hundreds of free books in PDF
- Read thousands of books online for free
- Explore our innovative research tools – try the "Research Dashboard" now!
- Sign up to be notified when new books are published
- Purchase printed books and selected PDF files

Thank you for downloading this PDF. If you have comments, questions or just want more information about the books published by the National Academies Press, you may contact our customer service department toll-free at 888-624-8373, visit us online, or send an email to feedback@nap.edu.

This book plus thousands more are available at http://www.nap.edu.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

5

Descriptions of Some Forensic Science Disciplines

This chapter describes the methods of some of the major forensic science disciplines. It focuses on those that are used most commonly for investigations and trials as well as on those that have been cause for concern in court or elsewhere because their reliability has not been sufficiently established in a systematic (scientific) manner in accordance with the principles discussed in Chapter 4. The chapter focuses primarily on the forensic science disciplines' capability for providing evidence that can be presented in court. As such, there is considerable discussion about the reliability and precision of results—attributes that factor into probative value and admissibility decisions. It should be recalled, however, that forensic science also provides great value to law enforcement investigations, and even those forensic science disciplines whose scientific foundation is currently limited might have the capacity (or the potential) to provide probative information to advance a criminal investigation.[1] This chapter also provides the committee's summary assessment of each of these disciplines.[2]

---

[1] For example, forensic odontology might not be sufficiently grounded in science to be admissible under *Daubert*, but this discipline might be able to reliably exclude a suspect, thereby enabling law enforcement to focus its efforts on other suspects. And forensic science methods that do not meet the standards of admissible evidence might still offer leads to advance an investigation.

[2] The chapter does not discuss eyewitness identification or line-ups, because these techniques do not normally rely on forensic scientists for analysis or implementation. They clearly are of major importance for investigations and trials, and their effective use and interpretation relies on scientific knowledge and continuing research. For similar reasons, this chapter does not delve into the polygraph. The validity of polygraph testing for security screening was addressed

Because forensic science aims to glean information from a wide variety of clues and evidence associated with a crime, it deals with a broad range of tools and with evidence of highly variable quality. In general, the forensic science disciplines are pragmatic, with practitioners adopting, adapting, or developing whatever tools and technological aids they can to distill useful information from crime scene evidence. Many forensic science methods have been developed in response to such evidence—combining experience-based knowledge with whatever relevant science base exists in order to create a procedure that returns useful information. Although some of the techniques used by the forensic science disciplines—such as DNA analysis, serology, forensic pathology, toxicology, chemical analysis, and digital and multimedia forensics—are built on solid bases of theory and research, many other techniques have been developed heuristically. That is, they are based on observation, experience, and reasoning without an underlying scientific theory, experiments designed to test the uncertainties and reliability of the method, or sufficient data that are collected and analyzed scientifically.

In the course of its deliberations, the committee received testimony from experts in many forensic science disciplines concerning current practices, validity, reliability and errors, standards, and research.[3] From this testimony and from many written submissions, as well as from the personal experiences of the committee members, the committee developed the consensus views presented in this chapter.

BIOLOGICAL EVIDENCE

Biological evidence is provided by specimens of a biological origin that are available in a forensic investigation. Such specimens may be found at the scene of a crime or on a person, clothing, or weapon. Some—for example, pet hairs, insects, seeds, or other botanical remnants—come from the crime scene or from an environment through which a victim or suspect has recently traversed. Other biological evidence comes from specimens obtained directly from the victim or suspect, such as blood, semen, saliva, vaginal secretions, sweat, epithelial cells, vomitus, feces, urine, hair, tissue, bones, and microbiological and viral agents. The most common types of biological evidence collected for examination are blood, semen, and saliva. Human biological evidence that contains nuclear DNA can be particularly valuable because the possibility exists to associate that evidence with one individual with a degree of reliability that is acceptable for criminal justice.

---

in National Research Council, Committee to Review the Scientific Evidence on the Polygraph. 2003. *The Polygraph and Lie Detection.* Washington, DC: The National Academies Press. It does not cover forensic pathology, because that field is addressed in Chapter 9.

[3] A complete list of those who provided testimony to the committee is included in Appendix B.

Sample Data and Collection

At the crime scene, biological evidence is located, documented, collected, and preserved for subsequent analysis in the crime laboratory. Locating and recognizing biological evidence can be more difficult than a layperson would presume. For example, blood is not always red, some red substances are not blood, and most biological evidence, such as saliva or semen, is not readily visible. Crime scene investigators locate biological evidence through tests that screen for the presence of a particular biological fluid (e.g., blood, semen, saliva), and investigators have a choice of techniques.[4] For blood they might use an alternate light source (ALS) at 415nm, the wavelength under which bloodstains absorb light and are thus more visible to the naked eye. Most commonly, though, the screening test for blood is a catalytic chemical test that turns color or luminesces in the presence of blood. Scene investigators may also use Luminol, fluorescein, or crystal violet to identify areas at the scene where attempts were made to clean a bloody crime scene.

These tests for blood may also locate other evidence that should be collected and taken to the laboratory for analysis. Recently, immunological tests that can identify human hemoglobin or glycophorin A have become available. These are blood-specific proteins that can be demonstrated to be of human origin. At some point in the future, these immunological tests may replace standard chemical tests, and, although more expensive, they are more specific because they identify blood conclusively instead of just presumptively. Investigators also have several techniques for locating semen at the crime scene. Commonly they rely on an ALS, under which semen, other biological fluids, and some other evidence will luminesce. More recently, immunological tests can be used to identify seminal plasma proteins, for example, prostate specific antigen (p30 or PSA) or semenogelin.[5]

Finding saliva at the scene is mostly happenstance. Although it luminesces with the ALS at specific wavelengths, the glow is not as strong, and a weaker ALS light source may not highlight it well and possibly not at all. Thus, it can be easily missed. Screening tests for saliva are chemical tests that identify amylase, an enzyme occurring in high concentrations in saliva. But the screening is not definitive, because other types of tissue also

---

[4] Interpreting the results of any screening test requires expertise and experience. Many crime scene investigators have the requisite experience, but they may lack a scientific background, and it is not always straightforward to correctly interpret the results of screening tests. Crime scene investigations that require science-based screening tools are most reliable if someone is involved who understands the physics and chemistry of those tools.

[5] I. Sato, M. Sagi, A. Ishiwari, H. Nishijima, E. Ito, and T. Mukai. 2002. Use of the "SMITEST" PSA card to identify the presence of prostate-specific antigen in semen and male urine. *Forensic Science International* 127(1-2):71-74.

contain amylase, including the particular type (AMY 1) that is associated with saliva.

Analyses

Although the forensic use of nuclear DNA is barely 20 years old, DNA typing is now universally recognized as the standard against which many other forensic individualization techniques are judged. DNA enjoys this preeminent position because of its reliability and the fact that, absent fraud or an error in labeling or handling, the probabilities of a false positive are quantifiable and often miniscule. However, even a very small (but nonzero) probability of false positive can affect the odds that a suspect is the source of a sample with a matching DNA profile.[6] The scientific bases and reliability of other types of biological analysis are also well established, but absent nuclear DNA, they can only narrow the field of suspects, not suggest any particular individual.

Testing biological evidence in the laboratory involves the use of a logical sequence of analyses designed to identify what a substance is and then from whom it came. The sequence begins with a forensic biologist locating the substance on the evidence. This is followed by a presumptive test that would give more information about the substance, typically using the same tests employed by scene investigators: the ALS, enzymatic, chemical, or immunological tests. Once the material (e.g., blood, semen, or saliva) is known, an immunological test or a human DNA test is run to determine whether the sample comes from a human or an animal.

The final step in the analytical sequence procedure is to identify the source of the biological material. If a sufficient sample is present and is probative, the forensic biologist prepares the material for DNA testing. The analyst who conducts the DNA test may or may not be the same person who examines the original physical evidence, depending on laboratory policies.

A decision might be required regarding the type of DNA testing to employ. Two primary types of DNA tests are conducted in U.S. forensic laboratories: nuclear testing and mitochondrial DNA (mtDNA) testing, with several variations of the former. For most biological evidence having evidentiary significance, forensic DNA laboratories employ nuclear testing routinely,[7] and testing for the 13 core Short Tandem Repeat (STR)

---

[6] W.C. Thompson, F. Taroni, and C.G.G. Aitken. 2003. How the probability of a false positive affects the value of DNA evidence. *Journal of Forensic Sciences* 48(1):47-54.

[7] T.R. Moretti, A.L. Baumstark, D.A. Defenbaugh, K.M. Keys, J.B. Smerick, and B. Budowle B. 2001. Validation of short tandem repeats (STRs) for forensic usage: Performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *Journal of Forensic Sciences* 46(3):647-660.

polymorphisms is the first line of attack.[8] The results are entered into the Federal Bureau of Investigation's (FBI's) Combined DNA Indexing System (CODIS) and are searched against DNA profiles already in one of three databases: a convicted felon database, a forensic database containing DNA profiles from crime scenes, and a database of DNA from unidentified persons.

Sometimes the evidence dictates testing just for Y STRs, which assesses only the Y (male) chromosome. In sexual assaults for which only small amounts of male nuclear DNA are available (e.g., a large excess of vaginal DNA), it is possible to obtain a Y STR profile of the male who left the semen. Unlike the 13 core loci used in CODIS searches, where a match of all 13 is a strong indicator that both samples come from the same individual, Y STR testing is not as definitive with respect to identifying a single person. A third nuclear test involves the analysis of single nucleotide polymorphisms (SNPs). Although no public forensic DNA laboratory in the United States is routinely analyzing forensic evidence for SNPs, the utility of this genomic information for cases in which the DNA is too damaged to allow standard testing has garnered attention since its use in the World Trade Center identification effort.[9]

If insufficient nuclear DNA is present for STR testing, or if the existing nuclear DNA is degraded, two options potentially are available. One technique amplifies the amount of DNA available, although this technique is not widely available in U.S. forensic laboratories. A second alternative is to sequence mitochondrial DNA (mtDNA). Since 1996, it has been possible to compare single-source crime scene samples and samples from the victim or defendant on the basis of mtDNA. Four FBI-supported mtDNA laboratories and a few private mtDNA laboratories conduct DNA casework. This technique has been particularly helpful with regard to hairs—which do not contain enough nuclear DNA to enable analysis with current methods unless the root is present—and bones and teeth. Because it measures only a single locus of the genome, mtDNA analysis is much less discriminating than nuclear DNA analysis; all people with a common female ancestor (within the past few generations) share a common profile. But mtDNA testing has forensic value in its ability to include or exclude an individual as its source.

Laboratories entering the results of forensic DNA testing into CODIS must meet specific quality guidelines, which include the requirement that

---

[8] Some laboratories are now using 16 loci, 13 of which are the original core loci.

[9] B. Leclair, R. Shaler, G.R. Carmody, K. Eliason, B.C.Hendrickson, T. Judkins, M.J. Norton, C. Sears, and T. Scholl. 2007. Bioinformatics and human identification in mass fatality incidents: The World Trade Center disaster. *Journal of Forensic Sciences* 52(4):806-819. Epub May 25, 2007.

the laboratory be accredited and that specific procedures be in place and followed. In accredited laboratories, forensic DNA personnel must take proficiency tests and must meet specific educational and training requirements. (See Chapter 8 for further discussion.) Laboratory analyses are conducted by scientists with degrees ranging from a bachelor's degree in science to a doctoral degree. Each forensic DNA laboratory has a technical leader, who normally must meet additional experience and educational requirements.

Although DNA laboratories are expected to conduct their examinations under stringent quality controlled environments, errors do occasionally occur. They usually involve situations in which interpretational ambiguities occur or in which samples were inappropriately processed and/or contaminated in the laboratory. Errors also can occur when there are limited amounts of DNA, which limits the amount of test information and increases the chance of misinterpretation. Casework reviews of mtDNA analysis suggest a wide range in the quality of testing results that include contamination, inexperience in interpreting mixtures, and differences in how a test is conducted.[10]

Reporting of Results

FBI quality guidelines require that reports from forensic DNA analysis must contain, at a minimum, a description of the evidence examined, a listing of the loci analyzed, a description of the methodology, results and/or conclusions, and an interpretative statement (either quantitative or qualitative) concerning the inference to be drawn from the analysis.[11]

---

[10] Personal communication, Terry Melton, Mitotyping Laboratory. December 2007. See also L. Prieto; A. Alonso; C. Alves; M. Crespillo; M. Montesino; A. Picornell; A. Brehm; J.L. Ramirez; M.R. Whittle; M.J. Anjos; I. Boschi; J. Buj; M. Cerezo; S. Cardoso; R. Cicarelli; D. Comas; D. Corach; C. Doutremepuich; R.M. Espinheira; I. Fernandez-Fernandez; S. Filippini; Julia Garcia-Hirschfeld; A. Gonzalez; B. Heinrichs; A. Hernandez; F.P.N. Leite; R.P. Lizarazo; A.M. Lopez-Parra; M. Lopez-Soto; J.A. Lorente; B. Mechoso; I. Navarro; S. Pagano; J.J. Pestano; J. Puente; E. Raimondi; A. Rodriguez-Quesada; M.F. Terra-Pinheiro; L. Vidal-Rioja; C. Vullo; A. Salas. 2008. GEP-ISFG collaborative exercise on mtDNA: Reflections about interpretation, artefacts and DNA mixtures. *Forensic Science International: Genetics* 2(2):126-133; and A. Salas, L. Prieto, M. Montesino, C. Albarrán, E. Arroyo, M. Paredes-Herrera, A. Di Lonardo, C. Doutremepuich, I. Fernández-Fernández, A. de la Vega. 2005. Mitochondrial DNA error prophylaxis: Assessing the causes of errors in the GEP'02-03 proficiency testing trial. *Forensic Science International* 148(2-3):191-198.

[11] DNA Advisory Board. 2000. Quality assurance standards for forensic DNA testing laboratories. *Forensic Science Communications* 2(3). Available at www.bioforensics.com/conference04/TWGDAM/Quality_Assurance_Standards_2.pdf.

Summary Assessment

Unlike many forensic techniques that were developed empirically within the forensic science community, with limited foundation in scientific theory or analysis, DNA analysis is a fortuitous by-product of cutting-edge science. Eminent scientists contributed their expertise to ensuring that DNA evidence offered in a courtroom would be valid and reliable (e.g., in the 1989 New York case, *People v. Castro*), and by 1996 the National Academy of Sciences had convened two committees that issued influential recommendations on handling DNA forensic science.[12] As a result, principles of statistics and population genetics that pertain to DNA evidence were clarified, the methods for conducting DNA analyses and declaring a match became less subjective, and quality assurance and quality control protocols were designed to improve laboratory performance.

DNA analysis is scientifically sound for several reasons: (1) there are biological explanations for individual-specific findings; (2) the 13 STR loci used to compare DNA samples were selected so that the chance of two different people matching on all of them would be extremely small; (3) the probabilities of false positives have been explored and quantified in some settings (even if only approximately); (4) the laboratory procedures are well specified and subject to validation and proficiency testing; and (5) there are clear and repeatable standards for analysis, interpretation, and reporting. DNA analysis also has been subjected to more scrutiny than any other forensic science discipline, with rigorous experimentation and validation performed prior to its use in forensic investigations. As a result of these characteristics, the probative power of DNA is high. Of course, DNA evidence is not available in every criminal investigation, and it is still subject to errors in handling that can invalidate the analysis. In such cases, other forensic techniques must be applied. The probative power of these other methods can be high, alone or in combination with other evidence. This power likely can be improved by strengthening the methods' scientific foundations and practice, as has occurred with forensic DNA analysis.

ANALYSIS OF CONTROLLED SUBSTANCES

The term "illicit drugs" is widely used to describe abused substances. Other terms that are used include "abused drugs," "illegal drugs," "street drugs," and, in the United States, "controlled substances." The latter term refers specifically to drugs that are controlled by federal and state laws.[13]

---

[12] National Research Council. 1992. *DNA Technology in Forensic Science*. Washington, DC: National Academy Press; National Research Council. 1996. *The Evaluation of Forensic DNA Evidence: An Update*. Washington, DC: National Academy Press.

[13] See, e.g., 21 U.S.C.A. § 802(6).

The analysis of controlled substances is a mature forensic science discipline and one of the areas with a strong scientific underpinning. The analytical methods used have been adopted from classical analytical chemistry, and there is broad agreement nationwide about best practices.[14] In 1997, the U.S. Drug Enforcement Administration and the Office of National Drug Control Policy co-sponsored the formation of the Technical Working Group for the Analysis of Seized Drugs, now known as the Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG). This organization brings together more than 20 forensic practitioners from all over the world to develop standards for the analysis and reporting of illicit drug cases. Their standards are being widely adopted by drug analysis laboratories in the United States and worldwide.

Sample Data and Collection

Controlled substances typically are seized by police officers, narcotics agents, and detectives through undercover buys, raids on drug houses and clandestine drug laboratories, and seizures on the streets. In some cases, forensic chemists are sent to clandestine laboratory operations to help render the laboratory safe and help with evidence collection. The seized drugs may be in the form of powders or adulterated powders, chunks of smokeable or injectable material, legitimate and clandestine tablets and capsules, or plant materials or plant extracts.

Analyses

Controlled substances are analyzed by well-accepted standard schemes or protocols. Few drug chemists have the requisite botanical background to identify any common illicit plants other than marijuana; thus, in cases that require botanical identification, the assistance of outside experts is enlisted.

Sampling can be a major issue in the analysis of controlled substances. Although sometimes only trace amounts of a drug are present (e.g., in a syringe used to inject heroin), at other times there are hundreds or thousands of packages of drugs or very large bags or bales. SWGDRUG and others have proposed statistical and nonstatistical methods for sampling,[15] and a wide variety of methods are used.

Most controlled substances are subjected first to a field test for pre-

---

[14] See F. Smith and J.A. Siegel (eds.). 2004. *Handbook of Forensic Drug Analysis*. Burlington, MA: Academic Press.

[15] Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) Recommendations. Available at www.swgdrug.org/approved.htm.

sumptive identification. This is followed by gas chromatography-mass spectrometry (GC-MS), in which chromatography separates the drug from any diluents or excipients, and then mass spectrometry is used to identify the drug. This is the near universal test for identifying unknown substances. Marijuana is an exception, because it is identified normally through a sequence of tests—a presumptive color test, followed by low-powered microscopic identification, and finally by thin-layer chromatography.

Reporting of Results

Most drug chemists produce terse reports for attorneys and courts. The reports contain administrative data and a short description of the evidence. The weight or number of exhibits is stated and then the results of the analysis. A typical report for a marijuana case might read as follows:

Received: Item 1—a sealed plastic bag containing 25.6 g of green-brown plant material.

Results: The green-brown plant material in item 1 was identified as marijuana.

Some laboratories might mention the tests that were conducted, but in most cases the spectra, chromatograms, and other evidence of the analysis and the chemist's notes are not submitted. Likewise, possible sources of error and statistical data are not commonly included. From a scientific perspective, this style of reporting is often inadequate, because it may not provide enough detail to enable a peer or other courtroom participant to understand and, if needed, question the sampling scheme, process(es) of analysis, or interpretation.

Summary Assessment

The chemical foundations for the analysis of controlled substances are sound, and there exists an adequate understanding of the uncertainties and potential errors. SWGDRUG has established a fairly complete set of recommended practices.[16] It also provides pointers to a number of guidelines for statistical sampling, both for illegal drugs per se (created by the European Network of Forensic Science Institutes) and for materials more generally (created by the American Society for Testing and Materials).

The SWGDRUG recommendations include a menu of analytical chemistry techniques that are considered acceptable in certain circumstances. Because this menu was constructed to be applicable worldwide, it includes

---

[16] See www.swgdrug.org/approved.htm.

options that allow laboratories to substitute a concatenation of simple methods if they do not have access to the preferred analytical equipment (e.g., GC-MS). It is questionable, however, whether all of the possible combinations recommended by SWGDRUG would be acceptable in a scientific sense, if one's goal were to identify and classify a completely unknown substance. The committee has been told that experienced forensic chemists and good forensic laboratories understand which tests (or combinations of tests) provide adequate reliability, but the SWGDRUG recommendations do not ensure that these tests will be used. This ambiguity would be a less significant issue if the reports presented in court contained sufficient detail about the methods of analysis.

FRICTION RIDGE ANALYSIS

Fingerprints, palm prints, and sole prints have been used to identify people for more than a century in the United States. Collectively, the analysis of these prints is known as "friction ridge analysis," which consists of experience-based comparisons of the impressions left by the ridge structures of volar (hands and feet) surfaces. Friction ridge analysis is an example of what the forensic science community uses as a method for assessing "individualization"—the conclusion that a piece of evidence (here, a pattern left by friction ridges) comes from a single unambiguous source. Friction ridge analysis shares similarities with other experience-based methods of pattern recognition, such as those for footwear and tire impressions, toolmarks, and handwriting analysis, all of which are discussed separately below.

Friction ridge analysis is performed in various settings, including accredited crime laboratories and nonaccredited facilities. Nonaccredited facilities may be crime laboratories, police "identification units," or private practice (consultants). In some instances, the latent print examiner is employed solely to perform latent print casework. Some examiners may also perform other types of forensic casework (e.g., footwear and tire impressions, firearms analysis). In some agencies, fingerprint examiners also are required to respond to crime scenes and can be sworn officers who also perform police officer/detective duties.

The training of personnel to perform latent print identifications varies from agency to agency. Agencies may have a formalized training program, may use an informal mentoring process, or may send new examiners to a one- to two-week course. The International Association for Identification (IAI) offers a training publication, "Friction Ridge Skin Identification Training Manual,"[17] and the Scientific Working Group on Friction Ridge

---

[17] International Association for Identification. *Friction Ridge Skin Identification Training Manual*. Available at www.theiai.org.

Analysis, Study and Technology (SWGFAST) offers a guideline, "Training to Competency for Latent Print Examiners."[18] Although these are excellent resources, they are not required, and there is no auditing of the content of training programs developed by nonaccredited agencies. The IAI also offers a certification test that measures both the knowledge and skill of latent print examiners; however, not all agencies require latent print examiners to achieve and maintain certification.

Method of Data Collection and Analysis

The technique used to examine prints made by friction ridge skin is described by the acronym ACE-V: "Analysis, Comparison, Evaluation, and Verification."[19] It has been described in forensic literature as a means of comparative analysis of evidence since 1959.[20] The process begins with the analysis of the unknown friction ridge print (now often a digital image of a latent print). Many factors affect the quality and quantity of detail in the latent print and also introduce variability in the resulting impression. The examiner must consider the following:

(1) Condition of the skin—natural ridge structure (robustness of the ridge structure), consequences of aging, superficial damage to the skin, permanent scars, skin diseases, and masking attempts.

(2) Type of residue—natural residue (sweat residue, oily residue, combinations of sweat and oil); other types of residue (blood, paint, etc.); amount of residue (heavy, medium, or light); and where the residue accumulates (top of the ridge, both edges of the ridge, one edge of the ridge, or in the furrows).

(3) Mechanics of touch—underlying structures of the hands and feet (bone creates areas of high pressure on the surface of the skin); flexibility of the ridges, furrows, and creases; the distance adjacent ridges can be pushed together or pulled apart during lateral movement; the distance the length of a ridge might be compressed or stretched; the rotation of ridge systems during torsion; and the effect of ridge flow on these factors.

(4) Nature of the surface touched—texture (rough or smooth), flexibility (rigid or pliable), shape (flat or curved), condition (clean or dirty), and background colors and patterns.

---

[18] SWGFAST. 2002. *Training to Competency for Latent Print Examiners.* Available at www. SWGFAST.org.

[19] Ashbaugh, op. cit.; Triplette and Cooney, op. cit.; J. Vanderkolk. 2004. ACE-V: A model. *Journal of Forensic Identification* 54(1):45-52; SWGFAST. 2002. *Friction Ridge Examination Methodology for Latent Print Examiners.* Available at www.SWGFAST.org.

[20] R.A. Huber. 1959-1960. Expert witness. *Criminal Law Quarterly* 2:276-296.

(5) Development technique—chemical signature of the technique and consistency of the chemical signature across the impression.

(6) Capture technique—photograph (digital or film) or lifting material (e.g., tape or gelatin lifter).

(7) Size of the latent print or the percentage of the surface that is available for comparison.

The examiner also must perform an analysis of the known prints (taken from a suspect or retrieved from a database of fingerprints), because many of the same factors that affect the quality of the latent print can also affect the known prints.

If the latent print does not have sufficient detail for either identification or exclusion, it does not undergo the remainder of the process (comparison and evaluation). These insufficient prints are often called "of no value" or "not suitable" for comparison. Poor-quality known prints also will end the examination. If the examiner deems that there is sufficient detail in the latent print (and the known prints), the comparison of the latent print to the known prints begins.

Visual comparison consists of discerning, visually "measuring," and comparing—within the comparable areas of the latent print and the known prints—the details that correspond. The amount of friction ridge detail available for this step depends on the clarity of the two impressions. The details observed might include the overall shape of the latent print, anatomical aspects, ridge flows, ridge counts, shape of the core, delta location and shape, lengths of the ridges, minutia location and type, thickness of the ridges and furrows, shapes of the ridges, pore position, crease patterns and shapes, scar shapes, and temporary feature shapes (e.g., a wart).

At the completion of the comparison, the examiner performs an evaluation of the agreement of the friction ridge formations in the two prints and evaluates the sufficiency of the detail present to establish an identification (source determination).[21] Source determination is made when the examiner concludes, based on his or her experience, that sufficient quantity and quality of friction ridge detail is in agreement between the latent print and the known print. Source exclusion is made when the process indicates sufficient disagreement between the latent print and known print. If neither an identification nor an exclusion can be reached, the result of the comparison is inconclusive. Verification occurs when another qualified examiner repeats the observations and comes to the same conclusion, although the second examiner may be aware of the conclusion of the first. A more complete de-

---

[21] Ashbaugh, op. cit.; SWGFAST. 2002. *Friction Ridge Examination Methodology for Latent Print Examiners.*

scription of the steps of ACE-V and an analysis of its limitations is provided in a paper by Haber and Haber.[22]

Although some Automated Fingerprint Identification Systems (AFIS) permit fully automated identification of fingerprint records related to criminal history (e.g., for screening job applicants), the assessment of latent prints from crime scenes is based largely on human interpretation. Note that the ACE-V method does not specify particular measurements or a standard test protocol, and examiners must make subjective assessments throughout. In the United States, the threshold for making a source identification is deliberately kept subjective, so that the examiner can take into account both the quantity and quality of comparable details. As a result, the outcome of a friction ridge analysis is not necessarily repeatable from examiner to examiner. In fact, recent research by Dror[23] has shown that experienced examiners do not necessarily agree with even their own past conclusions when the examination is presented in a different context some time later.

This subjectivity is intrinsic to friction ridge analysis, as can be seen when comparing it with DNA analysis. For the latter, 13 specific segments of DNA (generally) are compared for each of two DNA samples. Each of these segments consists of ordered sequences of the base pairs, called A, G, C, and T. Studies have been conducted to determine the range of variation in the sequence of base pairs at each of the 13 loci and also to determine how much variation exists in different populations. From these data, scientists can calculate the probability that two DNA samples from different people will have the same permutations at each of the 13 loci.

By contrast, before examining two fingerprints, one cannot say a priori which features should be compared. Features are selected during the comparison phase of ACE-V, when a fingerprint examiner identifies which features are common to the two impressions and are clear enough to be evaluated. Because a feature that was helpful during a previous comparison might not exist on these prints or might not have been captured in the latent impression, the process does not allow one to stipulate specific measurements in advance, as is done for a DNA analysis. Moreover, a small stretching of distance between two fingerprint features, or a twisting of angles, can result from either a difference between the fingers that left the prints or from distortions from the impression process. For these reasons, population statistics for fingerprints have not been developed, and friction ridge analysis relies on subjective judgments by the examiner. Little research

[22] L. Haber and R.N. Haber. 2008. Scientific validation of fingerprint evidence under *Daubert. Law, Probability, and Risk* 7(2):87-109.

[23] I.E. Dror and D. Charlton. 2006. Why experts make errors. *Journal of Forensic Identification* 56(4):600-616.

has been directed toward developing population statistics, although more would be feasible.[24]

Methods of Interpretation

The determination of an exclusion can be straightforward if the examiner finds detail in the latent print that does not match the corresponding part of the known print, although distortions or poor image quality can complicate this determination. But the criteria for identification are much harder to define, because they depend on an examiner's ability to discern patterns (possibly complex) among myriad features and on the examiner's experience judging the discriminatory value in those patterns. The clarity of the prints being compared is a major underlying factor. For 10-print fingerprint cards, which tend to have good clarity, even automated pattern-recognition software (which is not as capable as human examiners) is successful enough in retrieving matching sets from databases to enjoy widespread use. When dealing with a single latent print, however, the interpretation task becomes more challenging and relies more on the judgment of the examiner. The committee heard presentations from friction ridge experts who assured it that friction ridge identification works well when a careful examiner works with good-quality latent prints. Clearly, the reliability of the ACE-V process could be improved if specific measurement criteria were defined. Those criteria become increasingly important when working with latent prints that are smudged and incomplete, or when comparing impressions from two individuals whose prints are unusually similar.

The fingerprint community continues to assert that the ability to see latent print detail is an acquired skill attained only through repeated exposure to friction ridge impressions. In their view, a lengthy apprenticeship (typically two years, at the FBI Laboratory) with an experienced latent print examiner enables a new examiner to develop a sense of the rarity of features and groups of features; the rarity of particular kinds of ridge flows; the frequency of features in different areas of the hands and feet; the degree to which differences can be accounted for by mechanical distortion of the skin; a sense of how to extract detail from background noise; and a sense of how much friction ridge detail could be common to two prints from different

---

[24] See, e.g., E. Gutiérrez, V. Galera, J.M. Martínez, and C. Alonso. 2007. Biological variability of the minutiae in the fingerprints of a sample of the Spanish population. *Forensic Science International* 172(2-3):98-105. For information about the basic availability of data, see C. Champod, C.J. Lennard, P.A. Margot, and M. Stoilovic. 2004. *Fingerprints and other ridge skin impressions.* Boca Raton, FL: CRC Press; D.A. Stoney. 2001. "Measurement of Fingerprint Individuality." In: H.C. Lee and R.E. Gaensslen (eds.). *Advances in Fingerprint Technology.* 2nd ed. Boca Raton, FL: CRC Press; pp. 327-387.

sources.[25] From this base of experience, the fingerprint community asserts that the latent print examiner learns to judge whether there is sufficient detail (which varies with image quality) to make a source determination during the evaluation phase of ACE-V.

The latent print community in the United States has eschewed numerical scores and corresponding thresholds, because those developed to date[26] have been based only on minutia, not on the unique features of the friction ridge skin (e.g., lengths of ridges, shapes of ridges, crease lengths and shapes, scar lengths and shapes). Additionally, thresholds based on counting the number of features that correspond, lauded by some as being more "objective," are still based on primarily subjective criteria—an examiner must have the visual expertise to discern the features (most important in low-clarity prints) and must determine that they are indeed in agreement. A simple point count is insufficient for characterizing the detail present in a latent print; more nuanced criteria are needed, and, in fact, likely can be determined.

Reporting of Results

SWGFAST has promulgated three acceptable conclusions resulting from latent print comparison: individualization (or identification), exclusion, or inconclusive.[27] Although adherence to this standard is common, some latent print examiners report either "identification" or "negative" results. "Negative" (or sometimes "not identified") is an ambiguous conclusion, and it could mean excluded, inconclusive, or unable to locate after exhaustive search. It is problematic that the meaning of "negative" may be specific to a particular agency, examiner, or case.

Latent print examiners report an individualization when they are confident that two different sources could not have produced impressions with the same degree of agreement among details. This is a subjective assessment. There has been discussion regarding the use of statistics to assign match probabilities based on population distributions of certain friction ridge features. Current published statistical models, however, have not matured past counts of corresponding minutia and have not taken clarity into consideration. (This area is ripe for additional research.) As a result, the friction ridge community actively discourages its members from testifying in terms of the probability of a match; when a latent print examiner testifies that two

---

[25] T. Busey and J. Vanderkolk. 2005. Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research* 45:431-448.

[26] See, e.g., I.W. Evett and R.A. Williams. 1996. A review of the sixteen points fingerprint standard in England and Wales. *Journal of Forensic Identification* 46(1):49-73.

[27] SWGFAST. 2002. *Friction Ridge Examination Methodology for Latent Print Examiners.* Available at www.swgfast.org/Training_to_Competency_for_Latent_Print_Examiners_2.1.pdf.

impressions "match," they are communicating the notion that the prints could not possibly have come from two different individuals.

As noted in Chapter 3, Jennifer Mnookin of the University of California, Los Angeles School of Law summarized the reporting of fingerprint analyses as follows:

> At present, fingerprint examiners typically testify in the language of absolute certainty. Both the conceptual foundations and the professional norms of latent fingerprinting prohibit experts from testifying to identification unless they believe themselves certain that they have made a correct match. Experts therefore make only what they term 'positive' or 'absolute' identifications—essentially making the claim that they have matched the latent print to the one and only person in the entire world whose fingertip could have produced it . . . Given the general lack of validity testing for fingerprinting; the relative dearth of difficult proficiency tests; the lack of a statistically valid model of fingerprinting; and the lack of validated standards for declaring a match, such claims of absolute, certain confidence in identification are unjustified . . . Therefore, in order to pass scrutiny under Daubert, fingerprint identification experts should exhibit a greater degree of epistemological humility. Claims of 'absolute' and 'positive' identification should be replaced by more modest claims about the meaning and significance of a 'match.'[28]

Summary Assessment

Historically, friction ridge analysis has served as a valuable tool, both to identify the guilty and to exclude the innocent. Because of the amount of detail available in friction ridges, it seems plausible that a careful comparison of two impressions can accurately discern whether or not they had a common source. Although there is limited information about the accuracy and reliability of friction ridge analyses, claims that these analyses have zero error rates are not scientifically plausible.

ACE-V provides a broadly stated framework for conducting friction ridge analyses. However, this framework is not specific enough to qualify as a validated method for this type of analysis. ACE-V does not guard against bias; is too broad to ensure repeatability and transparency; and does not guarantee that two analysts following it will obtain the same results. For these reasons, merely following the steps of ACE-V does not imply that one is proceeding in a scientific manner or producing reliable results. A recent

---

[28] J.L. Mnookin. 2008. The validity of latent fingerprint identification: Confessions of a fingerprinting moderate. *Law, Probability and Risk* 7:127. See also the discussion in C. Champod. 2008. Fingerprint examination: Towards more transparency. *Law Probability and Risk* 7:111-118.

paper by Haber and Haber[29] presents a thorough analysis of the ACE-V method and its scientific validity. Their conclusion is unambiguous: "We have reviewed available scientific evidence of the validity of the ACE-V method and found none."[30] Further, they state:

> [W]e report a range of existing evidence that suggests that examiners differ at each stage of the method in the conclusions they reach. To the extent that they differ, some conclusions are invalid. We have analysed the ACE-V method itself, as it is described in the literature. We found that these descriptions differ, no single protocol has been officially accepted by the profession and the standards upon which the method's conclusions rest have not been specified quantitatively. As a consequence, at this time the validity of the ACE-V method cannot be tested.[31]

Recent legal challenges, *New Hampshire vs. Richard Langill*[32] and *Maryland vs. Bryan Rose*,[33] have also highlighted two important issues for the latent print community: documentation and error rate. Better documentation is needed of each step in the ACE-V process or its equivalent. At the very least, sufficient documentation is needed to reconstruct the analysis, if necessary. By documenting the relevant information gathered during the analysis, evaluation, and comparison of latent prints and the basis for the conclusion (identification, exclusion, or inconclusive), the examiner will create a transparent record of the method and thereby provide the courts with additional information on which to assess the reliability of the method for a specific case. Currently, there is no requirement for examiners to document which features within a latent print support their reasoning and conclusions.

Error rate is a much more difficult challenge. Errors can occur with any judgment-based method, especially when the factors that lead to the ultimate judgment are not documented. Some in the latent print community argue that the method itself, if followed correctly (i.e., by well-trained examiners properly using the method), has a zero error rate. Clearly, this assertion is unrealistic, and, moreover, it does not lead to a process of method improvement. The method, and the performance of those who use it, are inextricably linked, and both involve multiple sources of error (e.g., errors in executing the process steps, as well as errors in human judgment).

Some scientific evidence supports the presumption that friction ridge patterns are unique to each person and persist unchanged throughout a

---

[29] Mnookin, op. cit.

[30] Ibid., p. 19.

[31] Ibid.

[32] 157 N.H. 77, 945 A.2d 1 (N.H., April 04, 2008).

[33] No. K06-0545 (MD Cir. Ct. Oct. 19, 2007).

lifetime.[34] Uniqueness and persistence are necessary conditions for friction ridge identification to be feasible, but those conditions do not imply that anyone can reliably discern whether or not two friction ridge impressions were made by the same person. Uniqueness does not guarantee that prints from two different people are always sufficiently different that they cannot be confused, or that two impressions made by the same finger will also be sufficiently similar to be discerned as coming from the same source. The impression left by a given finger will differ every time, because of inevitable variations in pressure, which change the degree of contact between each part of the ridge structure and the impression medium. None of these variabilities—of features across a population of fingers or of repeated impressions left by the same finger—has been characterized, quantified, or compared.[35]

To properly underpin the process of friction ridge identification, additional research is also needed into ridge flow and crease pattern distributions on the hands and feet. This information could be used to limit the possible donor population of a particular print in a statistical approach and could provide examiners with a more robust understanding of the prevalence of different ridge flows and crease patterns. Additionally, more research is needed regarding the discriminating value of the various ridge formations and clusters of ridge formations.[36] This would provide examiners with a solid basis for the intuitive knowledge they have gained through experience and provide an excellent training tool. It also would lead to a good framework for future statistical models and provide the courts with additional information to consider when evaluating the reliability of the science. Recently, research has begun to build some of this basis.[37]

---

[34] F. Galton. 1892. *Fingerprints.* New York: MacMillan; H. Cummins and C. Midlo. 1943. *Finger Prints, Palms and Soles: An Introduction of Dermatoglyphics.* Philadelphia: The Blakiston Company; A. Hale. 1952. Morphogenesis of volar skin in the human fetus. *The American Journal of Anatomy* 91:147-173; S. Holt and L.S. Penrose. 1968. *The Genetics of Dermal Ridges.* Springfield, IL: Charles C Thomas Publishing; W. Montagna and P. Parakkal. 1974. *The Structure and Function of Skin.* New York: Academic Press; J. Raser and E. O'Shea. 2005. Noise in gene expression: Origins, consequences, and control. *Science* 39:2010-2013.

[35] Some in the friction ridge community point to an unpublished 1999 study by the Lockheed-Martin Corporation, the "50K vs. 50K Fingerprint Comparison Test," as evidence of the scientific validity of fingerprint "matchup." But that study has several major design and analysis flaws, as pointed out in D.H. Kaye. 2003. Questioning a courtroom proof of the uniqueness of fingerprints. *International Statistical Review* 71(3):524. Moreover, even if it were valid, the study provides only a highly optimistic estimate of the reliability of friction ridge analyses, biased toward highly favorable conditions.

[36] Haber and Haber also provide a sensible research agenda for enhancing the validity of fingerprint comparisons.

[37] E.g., C. Neumann, C. Champod, R. Puch-Solis, N. Egli, A. Anthonioz, and A. Bromage-Griffiths. 2007. Computation of likelihood ratios in fingerprint identification for configurations of any number of minutiae. *Journal of Forensic Sciences* 52(1):54-64; N.M. Egli,

There is also considerable room for research on the various factors that affect the quality of latent prints (e.g., condition of the skin, residue, mechanics of touch). Formal research could provide examiners with additional tools to support or refute distortion explanations. Currently, distortion and quality issues are typically based on "common sense" explanations or on information that is passed down through oral tradition from examiner to examiner. A criticism of the latent print community is that the examiners can too easily explain a "difference" as an "acceptable distortion" in order to make an identification.[38]

OTHER PATTERN/IMPRESSION EVIDENCE: SHOEPRINTS AND TIRE TRACKS

Other pattern evidence, also referred to as impression evidence, occurs when an object such as a shoe or a tire leaves an impression at the crime scene or on another object or a person. Impressions can be either two dimensional, such as shoeprints in dust, or three dimensional, such as tire track impressions in mud. Shoeprints and tire tracks are common types of impression evidence examined by forensic examiners, but the list of potential types of impression evidence is long. Examples include bite marks, markings on bullets and cartridge cases, ear prints, lip prints, toolmarks, some bloodstain patterns, and glove prints.[39] Although there are general approaches concerning the analytical sequence of various types of impression evidence, each has its own set of characteristics. For example, some types of impression evidence, such as those arising from footwear and tires, require knowledge of manufacturing and wear, while other types, such as ear prints and bloodstain patterns, do not. Because footwear and tire track impressions comprise the bulk of the examinations conducted, the remarks in this section are specifically focused on these analyses. Bite marks, markings on bullets and cartridge cases, and bloodstain patterns are covered in later sections in this chapter.

C. Champod, and P. Margot. 2007. Evidence evaluation in fingerprint comparison and automated fingerprint identification systems—Modelling within finger variability. *Forensic Science International* 167(2-3):189-195.

[38] U.S. Department of Justice, Office of the Inspector General. 2006. *A Review of the FBI's Handling of the Brandon Mayfield Case*. Office of the Inspector General Oversight and Review Division, January.

[39] M. Liukkonen, H. Majamaa, and J. Virtanen. 1996. The role and duties of the shoeprint/ toolmark examiner in forensic laboratories. *Forensic Science International* 82:99-108.

Sample Data and Collection

Impression evidence at the scene is generally of two types: latent (invisible to the naked eye) or patent (visible). The quality of impression evidence left at the scene cannot be controlled, but failures in the initial scene work used to collect, preserve, and possibly enhance the evidence will degrade the quality of the evidence eventually used for comparative analysis. After documentation at the scene, the evidence is preserved and possibly enhanced using techniques such as those based on chemistry (e.g., metal detection), physical characteristics (e.g., super glue fuming, powder dusting, casting), or transfer onto a contrasting surface (e.g., electrostatic transfer or gel lifting). The quality of the enhanced impression that is used for comparison will depend largely on the experience, training, and scientific knowledge of the scene investigator as well as the agency's resources.

Although some analysis of impression evidence might begin at the scene, the comparison of scene evidence to known exemplars occurs in the laboratory. The educational background of forensic scientists who examine shoeprints and tire track impressions runs the gamut from a high school diploma to scientists with Ph.D.s. Identifications are largely subjective and are based on the examiner's experience and on the number of individual, identifying characteristics in common with a known standard.

Analyses

The goal of impression evidence analysis is to identify a specific source of the impression, and the analytical process that this follows generally is an accepted sequence: identifying the class (group) characteristics of the evidence, followed by locating and comparing individual, identifying (also termed accidental or random) characteristics.[40]

Class characteristics of footwear and tires result from repetitive, controlled processes that are typically mechanical, such as those used to manufacture items in quantity. Although defined similarly by various authors, Bodziak describes footwear class characteristics as "an intentional or unavoidable characteristic that repeats during the manufacturing process and is shared by one or more other shoes."[41] For tires, Nause defines class characteristics as, "[p]hysical characteristics acquired during the manufacturing process (made from the same mold) that tires have in common."[42] He continues, "Class characteristics can often be combined to limit a tire impression to a very select group within the overall group bearing similar class

---

[40] Ibid.

[41] W.J. Bodziak. 1999. *Footwear Impression Evidence–Detection, Recovery, and Examination.* Boca Raton, FL: CRC Press, 2nd ed., p. 329.

[42] Nause, op. cit.

characteristics. (In the field of forensic tire evidence, class characteristics often refer to such things as design, pattern, size, shape, mold variations, etc.)."[43] Regardless of the type of impression evidence, class characteristics are not sufficient to conclude that any one particular shoe or tire made the impression. That latter step—which is not always possible—requires comparison of the individual identifying characteristics on the impression evidence with those on a shoe or tire that is suspected of leaving the impression. These individual characteristics occur during the normal use of an item, sometimes called wear and tear,[44] and are created by "random, uncontrolled processes."[45] For footwear, Bodziak writes that "individual identifying characteristics are characteristics that result when something is randomly added to or taken away from a shoe outsole that either causes or contributes to making that shoe outsole unique."[46] Such characteristics might include cuts, scratches, gouges, holes, or random inclusions that result from manufacturing, such as bubbles, and those that result from adherent substances, such as rocks, chewing gum, papers, or twigs.

Following analysis of the impression, an identification is determined or ruled out according to the number of individual characteristics the evidence has in common with the suspected source. But there is no defined threshold that must be surpassed, nor are there any studies that associate the number of matching characteristics with the probability that the impressions were made by a common source. It is generally accepted that the specific number of characteristics needed to assign a definite positive identification depends on the quality and quantity of these accidental characteristics and the criteria established by individual laboratories.[47] According to Cassidy, many factors and accidental characteristics are required before a positive identification can be established; however, the most important are the examiner's experience, the clarity of the impression, and the uniqueness of the characteristic.[48] Proficiency testing for examiners of impression evidence is available through Collaborative Testing Service, Inc., but the proficiency tests for footwear impressions include samples that are either a match or not a match[49]—that is, none of the samples included in the tests have the sort of ambiguities that would lead an experienced examiner to an "inconclusive"

---

[43] Ibid.

[44] M.J. Cassidy. 1980. *Footwear Identification.* Quebec, Canada: Government Printing Office Centre.

[45] K. Inman and N. Rudin. 2001. *Principles and Practice of Criminalistics.* Boca Raton, FL: CRC Press, p. 129.

[46] Ibid., p. 335.

[47] Liukkonen, Majamaa, and Virtanen, op. cit.

[48] Cassidy, op. cit.

[49] H. Majamaa and Y. Anja. 1996. Survey of the conclusions drawn of similar footwear cases in various crime laboratories. *Forensic Science International* 82:109-120.

conclusion. IAI has a certification program for footwear and tire track examiners.[50] The group's recommended course of study has 13 segments, and each segment includes a suggested reading list and practical and/or written exercises. The student must pass an examination. This course of study does not require an understanding of the scientific basis of the examinations, and it does not recommend the use of a scientific method. Also, there is no provision or recommendation for proficiency testing or continuing education. SWGTREAD, a group of footwear and tire track examiners formed by the FBI, recommends that a trainee candidate have (1) a bachelor's degree (preferably in a physical or natural science) from an accredited college or university; or (2) an associate degree or 60 college semester hours, plus two years of job-related forensic experience; or (3) a high school diploma or equivalent, plus four years of job-related forensic experience.[51]

Scientific Interpretation and Reporting of Results

For footwear evidence, Fawcett[52] and Bodziak[53] have attempted to assign probabilistic or statistical significance to impression comparisons. Generally, shoeprint and tire track examiners prefer nonstatistical language to report or to testify to the result of their findings. Terms such as "positive identification" and "nonidentification" can be used to indicate an identification or nonidentification, respectively, and "nonconclusive" would indicate situations in which the analysis falls short of either of the other two.[54]

In a European survey, examiners were given identical mock cases. Accidental, identifying characteristics were purposely put onto the sole of new shoes, and examiners were asked to make a statement concerning the strength of matches. The results of the survey concluded that there were considerable differences in the conclusions reached by different laboratories examining identical cases."[55] SWGTREAD recommends terminology such as:

- "identification" (definite conclusion of identity)
- "probably made" (very high degree of association)

---

[50] Recommended Course of Study for Footwear & Tire Track Examiners. 1995. Mendota Heights, MN: International Association for Identification.

[51] SWGTread. 2006. *Guide for Minimum Qualifications and Training for a Forensic Footwear and/or Tire Tread Examiner.* Available at www.theiai.org/guidelines/swgtread/qualifications_final.pdf.

[52] A.S. Fawcett. 1970. The role of the footmark examiner. *Journal of the Forensic Science Society* 10:227-244.

[53] Bodziak, op. cit., pp. 342-346.

[54] Ibid.

[55] H. Majamaa and Y. Anja., op. cit.

- "could have made" (significant association of multiple class characteristics)
- "inconclusive" (limited association of some characteristics)
- "probably did not make" (very high degree of nonassociation)
- "elimination" (definite exclusion)
- "unsuitable" (lacks sufficient detail for a meaningful comparison).

Additionally, SWGTREAD discourages the use of once common terminology, such as "consistent with" (acceptable when used to describe a similarity of characteristics), "match/no match," "responsible for/not responsible for," and "caused with/not caused with."[56] Neither the IAI nor SWGTREAD address the statistical evaluation of impression evidence.

Summary Assessment

The scientific basis for the evaluation of impression evidence is that mass-produced items (e.g., shoes, tires) pick up features of wear that, over time, individualize them. However, because these features continue to change as they are worn, elapsed time after a crime can undercut the forensic scientist's certainty. At the least, class characteristics can be identified, and with sufficiently distinctive patterns of wear, one might hope for specific individualization. However, there is no consensus regarding the number of individual characteristics needed to make a positive identification, and the committee is not aware of any data about the variability of class or individual characteristics or about the validity or reliability of the method. Without such population studies, it is impossible to assess the number of characteristics that must match in order to have any particular degree of confidence about the source of the impression.

Experts in impression evidence will argue that they accumulate a sense of those probabilities through experience, which may be true. However, it is difficult to avoid biases in experience-based judgments, especially in the absence of a feedback mechanism to correct an erroneous judgment. These problems are exacerbated with the less common types of impression evidence. For example, a European survey found that 42 laboratories conducted 28,093 shoeprint examinations and 41 laboratories conducted 591 tire track examinations, but only 14 laboratories conducted a total of 21 lip print examinations and 17 laboratories conducted a total of 100 ear print examinations.[57] Although one might argue that those who perform the

---

[56] SWGTREAD. 2006. *Standard Terminology for Expressing Conclusions of Forensic Footwear and Tire Impression Examinations.* Available at www.theiai.org/guidelines/swgtread/terminology_final.pdf.

[57] Liukkonen, Majamaa, and Virtanen, op. cit.

work in laboratories that conduct hundreds or thousands of evaluations of impression evidence develop useful experience and judgment, it is difficult to assert that the field has enough collective judgment about the variabilities in lip prints and ear prints based on tens of examinations. The community simply does not have enough data about the natural variability of those less frequent impressions, absent the presence of a clear deformity or scar, to infer whether the observed degree of similarity is significant.

Most of the research in the field is conducted in forensic laboratories, with the results published in trade journals, such as the *Journal of Forensic Identification*. With regard to reporting, SWGTREAD is moving toward the use of standard language to convey the conclusions reached.[58] But neither IAI nor SWGTREAD addresses the issue of what critical research should be done or by whom, critical questions that should be addressed include the persistence of individual characteristics, the rarity of certain characteristic types, and the appropriate statistical standards to apply to the significance of individual characteristics. Also, little if any research has been done to address rare impression evidence. Much more research on these matters is needed.

TOOLMARK AND FIREARMS IDENTIFICATION

Toolmarks are generated when a hard object (tool) comes into contact with a relatively softer object. Such toolmarks may occur in the commission of a crime when an instrument such as a screwdriver, crowbar, or wire cutter is used or when the internal parts of a firearm make contact with the brass and lead that comprise ammunition. The marks left by an implement such as a screwdriver or a firearm's firing pin depend largely on the manufacturing processes—and manufacturing tools—used to create or shape it, although other surface features (e.g., chips, gouges) might be introduced through post-manufacturing wear. Manufacturing tools experience wear and abrasion as they cut, scrape, and otherwise shape metal, giving rise to the theory that any two manufactured products—even those produced consecutively with the same manufacturing tools—will bear microscopically different marks. Firearms and toolmark examiners believe that toolmarks may be traced to the physical heterogeneities of an individual tool—that is, that "individual characteristics" of toolmarks may be uniquely associated with a specific tool or firearm and are reproduced by the use of that tool and only that tool.

The manufacture and use of firearms produces an extensive set of

---

[58] SWGTREAD. 2006. *Standard Terminology for Expressing Conclusions of Forensic Footwear and Tire Impression Examinations.* Available at www.theiai.org/guidelines/swgtread/terminology_final.pdf.

specialized toolmarks. Gun barrels typically are rifled to improve accuracy, meaning that spiral grooves are cut into the barrel's interior. The process of cutting these grooves into the barrel leaves marks and scrapes on the relatively softer metal of the barrel.[59] In turn, these markings are transferred to the softer metal of a bullet as it exits the barrel. Over time, with repeated use (and metal-to-metal scraping), the marks on a barrel (and the corresponding "stria" imparted to bullets) may change as individual imperfections are formed or as cleanliness of the barrel changes. The brass exterior of cartridge cases receive analogous toolmarks during the process of gun firing: the firing pin dents the soft primer surface at the base of the cartridge to commence firing, the primer area is forced backward by the buildup of gas pressure (so that the texture of the gun's breech face is impressed on the cartridge), and extractors and ejectors leave marks as they expel used cartridges and cycle in new ammunition.

Firearms examination is one of the more common functions of crime laboratories. Even small laboratories with limited services often perform firearms analysis. In addition to the analysis of marks on bullets and cartridges, firearms examination also includes the determination of the firing distance, the operability of a weapon, and sometimes the analysis of primer residue to determine whether someone recently handled a weapon. These broader aspects are not covered here.

Sample and Data Collection

When a tool is used in a crime, the object that contains the tool marks is recovered when possible. If a toolmark cannot be recovered, it can be photographed and cast. Test marks made by recovered tools can be made in a laboratory and compared with crime scene toolmarks.

In the early 1990s, the FBI and the Bureau of Alcohol, Tobacco, Firearms, and Explosives (ATF) developed separate databases of images of bullet and cartridge case markings, which could be queried to suggest possible matches. In 1996, the National Institute of Standards and Technology (NIST) developed data exchange standards that permitted the integration of the FBI's DRUGFIRE database (cartridge case images) and the ATF's CEASEFIRE database (then limited to bullet images). The current National Integrated Ballistic Information Network (NIBIN) includes images from both cartridge cases and bullets that are associated with crime scenes and is maintained by the ATF.

Periodically—and particularly in the wake of the Washington, DC,

---

[59] Although the metal and initial rifling are very similar, the cutting of the individual barrels, the finishing machining, and the cleaning and polishing begin the process of differentiation of the two sequentially manufactured barrels.

sniper attacks in 2002—the question has been raised of expanding the scope of databases like NIBIN to include images from test firings of newly manufactured firearms. In concept, this would permit downstream investigators who recover a cartridge case or bullet at a crime scene to identify the likely source firearm. Though two states (Maryland and New York) instituted such reference ballistic image databases for newly manufactured firearms, proposals to create such a database at the national level did not make substantial progress in Congress. A recent report of the National Academies, *Ballistic Imaging,* examined this option in great detail and concluded that "[a] national reference ballistic image database of all new and imported guns is not advisable at this time."[60]

Analyses

In both firearm and toolmark identification, it is useful to distinguish several types of characteristics that are considered by examiners. "Class characteristics" are distinctive features that are shared by many items of the same type. For example, the width of the head of a screwdriver or the pattern of serrations in the blade of a knife may be class characteristics that are common to all screwdrivers or knives of a particular manufacturer and/or model. Similarly, the number of grooves cut into the barrel of a firearm and the direction of "twist" in those grooves are class characteristics that can filter and restrict the range of firearms that match evidence found at a crime scene. "Individual characteristics" are the fine microscopic markings and textures that are said to be unique to an individual tool or firearm. Between these two extremes are "subclass characteristics" that may be common to a small group of firearms and that are produced by the manufacturing process, such as when a worn or dull tool is used to cut barrel rifling.

Bullets and cartridge cases are first examined to determine which class characteristics are present. If these differ from a comparison bullet or cartridge, further examination may be unnecessary. The microscopic markings on bullets and cartridge cases and on toolmarks are then examined under a comparison microscope (made from two compound microscopes joined by a comparison bridge that allows viewing of two objects at the same time). The unknown and known bullet or cartridge case or toolmark surfaces are compared visually by a firearms examiner, who can evaluate whether a match exists.

---

[60] National Research Council. 2008. *Ballistic Imaging.* Washington, DC: The National Academies Press, p. 5.

Scientific Interpretation

The task of the firearms and toolmark examiner is to identify the individual characteristics of microscopic toolmarks apart from class and subclass characteristics and then to assess the extent of agreement in individual characteristics in the two sets of toolmarks to permit the identification of an individual tool or firearm.

Guidance from the Association of Firearm and Tool Mark Examiners (AFTE)[61] indicates that an examiner may offer an opinion that a specific tool or firearm was the source of a specific set of toolmarks or a particular bullet striation pattern when "sufficient agreement" exists in the pattern of two sets of marks. The standards then define agreement as significant "when it exceeds the best agreement demonstrated between tool marks known to have been produced by different tools and is consistent with the agreement demonstrated by tool marks known to have been produced by the same tool."[62]

Knowing the extent of agreement in marks made by different tools, and the extent of variation in marks made by the same tool, is a challenging task. AFTE standards acknowledge that these decisions involve subjective qualitative judgments by examiners and that the accuracy of examiners' assessments is highly dependent on their skill and training. In earlier years, toolmark examiners relied on their past casework to provide a foundation for distinguishing between individual, class, and subclass characteristics. More recently, extensive training programs using known samples have expanded the knowledge base of examiners.

The emergence of ballistic imaging technology and databases such as NIBIN assist examiners in finding possible candidate matches between pieces of evidence, including crime scene exhibits held in other geographic locations. However, it is important to note that the final determination of a match is always done through direct physical comparison of the evidence by a firearms examiner, not the computer analysis of images. The growth of these databases also permits examiners to become more familiar with similarities in striation patterns made by different firearms. Newer imaging techniques assess toolmarks using three-dimensional surface measurement data, taking into account the depth of the marks. But even with more training and experience using newer techniques, the decision of the toolmark examiner remains a subjective decision based on unarticulated

---

[61] Theory of identification, range of striae comparison reports and modified glossary definitions—An AFTE Criteria for Identification Committee report. 1992. *Journal of the Association of Firearm and Tool Mark Examiners* 24:336-340.

[62] Ibid., p. 336.

standards and no statistical foundation for estimation of error rates.[63] The National Academies report, *Ballistic Imaging,* while not claiming to be a definitive study on firearms identification, observed that, "The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated." That study recognized the logic involved in trying to compare firearms-related toolmarks by noting that, "Although they are subject to numerous sources of variability, firearms-related toolmarks are not completely random and volatile; one can find similar marks on bullets and cartridge cases from the same gun," but it cautioned that, "A significant amount of research would be needed to scientifically determine the degree to which firearms-related toolmarks are unique or even to quantitatively characterize the probability of uniqueness."[64]

Summary Assessment

Toolmark and firearms analysis suffers from the same limitations discussed above for impression evidence. Because not enough is known about the variabilities among individual tools and guns, we are not able to specify how many points of similarity are necessary for a given level of confidence in the result. Sufficient studies have not been done to understand the reliability and repeatability of the methods. The committee agrees that class characteristics are helpful in narrowing the pool of tools that may have left a distinctive mark. Individual patterns from manufacture or from wear might, in some cases, be distinctive enough to suggest one particular source, but additional studies should be performed to make the process of individualization more precise and repeatable.

---

[63] Recent research has attempted to develop a statistical foundation for assessing the likelihood that more than one tool could have made specific marks by assessing consecutive matching striae, but this approach is used in a minority of cases. See A.A. Biasotti. 1959. A statistical study of the individual characteristics of fired bullets. *Journal of Forensic Sciences* 4:34; A.A. Biasotti and J. Murdock. 1984. "Criteria for identification" or "state of the art" of firearms and tool marks identification. *Journal of the Association of Firearms and Tool Mark Examiners* 16(4):16; J. Miller and M.M. McLean. 1998. Criteria for identification of tool marks. *Journal of the Association of Firearms and Tool Mark Examiners* 30(1):15; J.J. Masson. 1997. Confidence level variations in firearms identification through computerized technology. *Journal of the Association of Firearms and Tool Mark Examiners* 29(1):42. For a critique of this area and a comparison of scientific issues involving toolmark evidence and DNA evidence, see A. Schwartz. 2004-2005. A systemic challenge to the reliability and admissibility of firearms and tool marks identification. *Columbia Science and Technology Law Review* 6:2. For a rebuttal to this critique, see R.G. Nichols. 2007. Defending the scientific foundations of the firearms and tool mark identification discipline: Responding to recent challenges. *Journal of Forensic Sciences* 52(3):586-594.

[64] All quotes from National Research Council. 2008. *Ballistic Imaging.* Washington, DC: The National Academies Press, p. 3.

A fundamental problem with toolmark and firearms analysis is the lack of a precisely defined process. As noted above, AFTE has adopted a theory of identification, but it does not provide a specific protocol. It says that an examiner may offer an opinion that a specific tool or firearm was the source of a specific set of toolmarks or a bullet striation pattern when "sufficient agreement" exists in the pattern of two sets of marks. It defines agreement as significant "when it exceeds the best agreement demonstrated between tool marks known to have been produced by different tools and is consistent with the agreement demonstrated by tool marks known to have been produced by the same tool." The meaning of "exceeds the best agreement" and "consistent with" are not specified, and the examiner is expected to draw on his or her own experience. This AFTE document, which is the best guidance available for the field of toolmark identification, does not even consider, let alone address, questions regarding variability, reliability, repeatability, or the number of correlations needed to achieve a given degree of confidence.

Although some studies have been performed on the degree of similarity that can be found between marks made by different tools and the variability in marks made by an individual tool, the scientific knowledge base for toolmark and firearms analysis is fairly limited. For example, a report from Hamby, Brundage, and Thorpe[65] includes capsule summaries of 68 toolmark and firearms studies. But the capsule summaries suggest a heavy reliance on the subjective findings of examiners rather than on the rigorous quantification and analysis of sources of variability. Overall, the process for toolmark and firearms comparisons lacks the specificity of the protocols for, say, 13 STR DNA analysis. This is not to say that toolmark analysis needs to be as objective as DNA analysis in order to provide value. And, as was the case for friction ridge analysis and in contrast to the case for DNA analysis, the specific features to be examined and compared between toolmarks cannot be stipulated a priori. But the protocols for DNA analysis do represent a precisely specified, and scientifically justified, series of steps that lead to results with well-characterized confidence limits, and that is the goal for all the methods of forensic science.

ANALYSIS OF HAIR EVIDENCE

The basis for hair analyses as forensic evidence stems from the fact that human and animal hairs routinely are shed and thus are capable of being

---

[65] J.E. Hamby, D.J. Brundage, and J.W. Thorpe. 2009. The identification of bullets fired from 10 consecutively rifled 9mm Ruger pistol barrels—A research project involving 468 participants from 19 countries. Available online at http://www.fti-ibis.com/DOWNLOADS/Publications/10%20Barrel%20Article-%20a.pdf.

transferred from an individual to the crime scene, and from the crime scene to an individual. Forensic hair examiners generally recognize that various physical characteristics of hairs can be identified and are sufficiently different among individuals that they can be useful in including, or excluding, certain persons from the pool of possible sources of the hair. The results of analyses from hair comparisons typically are accepted as class associations; that is, a conclusion of a "match" means only that the hair could have come from any person whose hair exhibited—within some levels of measurement uncertainties—the same microscopic characteristics, but it cannot uniquely identify one person. However, this information might be sufficiently useful to "narrow the pool" by excluding certain persons as sources of the hair.

Although animal hairs might provide useful evidence in certain cases (e.g., animal poaching), animal hair analysis often can lead to an identification of only the type of animal, not the specific breed[66]; consequently, most (90 to 95 percent) of hair analyses refer to analyses of human hair. Human hairs from different parts of the body have different characteristics; Houck cautions strongly against drawing conclusions about hairs from one part of the body based on analyses of hairs from a different body part.[67]

Houck and Bisbing recommend as minimal training for hair examiners a bachelor's degree in a natural or applied science (e.g., chemistry, biology, forensic science), on-the-job training programs, and an annual proficiency test.[68]

Sample Data and Collection

Sample hairs received for analysis initially are examined macroscopically for certain broad features such as color, shaft form (e.g., straight, wavy, curved, kinked), length, and overall shaft thickness (e.g., fine, medium, coarse).

In the second stage of analysis, hairs are mounted on microscopic slides using a mounting medium that has the same refractive index (about 1.54) as the hair, to better view the microscopic features (see next section). One hair or multiple hairs from the same source may be mounted on a glass microscope slide with an appropriate cover slip, as long as each mounted hair is clearly visible. It is most important that questioned and known hairs are mounted in the same type of mounting medium.

During this examination, the hair analyst attempts to identify the part of the body from which the hair might have come, based on certain de-

---

[66] P.D. Barnett and R.R. Ogle. 1982. Probabilities and human hair comparison. *Journal of Forensic Sciences* 27(2):272-278.

[67] M.M. Houck and R.E. Bisbing. 2005. Forensic human hair examination and comparison in the 21st century. *Forensic Science Review* 17(1):7.

[68] Ibid., p. 12.

finable characteristics that distinguish hairs from various body locations. Occasionally, suspects can be eliminated on the basis of these simple microscopic characteristics.

A "control" or "comparison" group of hairs must be collected from a known hair source. A known head hair sample should consist of hairs from the five different areas of the scalp (top, front, back including nape, and both sides). Known hair samples should be obtained by a combination of pulling and combing from the sampled region. Ideally, a total of 50 hairs should be obtained from the scalp. A known pubic hair sample or a sample from any other somatic region should ideally consist of 25 hairs obtained by pulling and combing from different regions. A comparison can still be performed with less than the recommended number of hairs, but this may increase the likelihood of a false exclusion.[69]

Features from human hair analyses can be divided broadly into "major characteristics" and "secondary characteristics." The former category includes features such as color, treatment (e.g., dyed, bleached, curled, permed), pigment aggregation (e.g., streaked, clumped, patchy), and shaft form (e.g., wavy, straight, curly). Other major characteristics may include pigment distribution (e.g., uniform, peripheral, clustered), medulla appearance, if present (e.g., continuous, interrupted, or fragmented—and opaque or translucent), hair diameter, medullary index, and presence or absence of cortical fusi (e.g., root or shaft). Secondary characteristics include cuticular margin (e.g., smooth, serrated, looped, or cracked), pigment density (e.g., absent, sparse, heavy), pigment size (e.g., absent, fine, coarse), tip shape (e.g., tapered, cut, rounded, frayed, split), and shaft diameter (e.g., narrow or wide).[70]

Studies of Accuracy in Identification

In 1974, investigators Gaudette and Keeping described a system of hair analysis and used it in a study of pairwise comparisons among 861 hairs from 100 different persons.[71] They acknowledged that "the hair samples were not chosen from the population at random, but were selected so that the probability of two hairs being similar would be greater, if anything, than in the population at large."[72] From their assignment of probabilities, the authors estimated that the chance of asserting a difference between two

---

[69] Scientific Working Group on Materials Analysis (SWGMAT). 2005. Forensic human hair examination guidelines. *Forensic Science Communications* 7(2). Available at www.fbi. gov/hq/lab/fsc/backissu/april2005/standards/2005_04_standards02.htm.

[70] Ibid.

[71] B.D. Gaudette and E.S. Keeping. 1974. An attempt at determining probabilities in human scalp hair comparison. *Journal of Forensic Sciences* 19(3):599-606.

[72] Ibid., p. 65.

hairs from the same person is small, about 1 in 4,500.[73] This assignment of probabilities has since been shown to be unreliable.[74] Moreover, the study does not confirm the chance of asserting a match between two dissimilar hairs, and the authors acknowledge that, "due to the fact that so many of the characteristics coded are subjective—for example, color, texture—it was not possible to get complete reproducibility between two or more examiners coding the same hair."[75]

Barnett and Ogle raised four concerns with the Gaudette and Keeping study: (1) it relied on idealized (not from real life) test scenarios; (2) there was no objective basis for selecting the features; (3) the statistical analysis of data from the study was questionable; and (4) there was a possible examiner bias.[76] Gaudette attempted to address these concerns through a further study. However, this additional study involved only three hair examiners, in addition to the author. The author concluded that:

> . . . whereas hair is not generally a basis for positive personal identification, the presence of abnormalities or unusual features or the presence of a large number of different unknown hairs all similar to the standard can lead to a more positive conclusion. The problem, at present, lies in finding suitable additional characteristics [of hair, for effecting individualization]. Although there is basic agreement as to the value of the macroscopic and microscopic characteristics used, other characteristics are either unreliable or controversial. Physical characteristics such as refractive index, density, scale counts, tensile strength, and electrical properties have been proposed by some workers but have been attacked by others, and the general consensus is that they are of little use in hair comparison.[77]

In 1990, Wickenheiser and Hepworth attempted a study to address examiner bias in a small study with only two examiners. They reported that "no incorrect associations were made by either examiner."[78] But a study with only two examiners cannot offer accurate and precise estimates of bias in the population of examiners.

An attempt at an objective system for identifying "matches" among hair samples is presented in Verma et al., based on a neural network.[79]

---

[73] A later study on human pubic hairs (Caucasian only) estimated this probability as "about 1 in 800." B.D. Gaudette. 1976. Probabilities and human pubic hair comparisons. *Journal of Forensic Sciences* 21(3):514-517.

[74] P.D. Barnett and R.R. Ogle. 1982. Probabilities and human hair comparison. *Journal of Forensic Sciences* 27(2):272-278.

[75] Gaudette and Keeping, op. cit.

[76] Barnett and Ogle, op. cit.

[77] B.D. Gaudette. 1978. Some further thoughts on probabilities and human hair comparisons. *Journal of Forensic Sciences* 23(4):758-763, pp. 761-762.

[78] Wickenheiser and Hepworth, op. cit., p. 1327.

[79] M.S. Verma, L. Pratt, C. Ganesh, and C. Medina. 2002. Hair-MAP: A prototype automated system for forensic hair comparison and analysis. *Forensic Science International* 129:168-186.

According to the authors of this article, "The system accurately judged whether two populations of hairs came from the same person or from different persons 83 percent of the time."[80] The article states that 83 percent was obtained by testing the neural network on all possible pairs among 9 samples of hairs from 9 people (i.e., 81 combinations, of which 9 are "true matches" and 72 are "true mismatches"). Their *Table 3*[81] can be summarized as follows:

	System said "same"	System said "different"	
Same person	5	4	Total= 9
Different persons	9	64	Total=73

Because the total of these 4 numbers is 82, not 81, one presumes a typographical error in the table; as stated, the number of correct calls is (5 + 64)/81=0.85, or 85 percent. (If one of the counts, 5 or 64, is off by 1, the percentage would be 84 percent.) However, the table also shows that the neural network claimed 9 of the 73 different pairs as "same," for a false positive rate of 9/73=12 percent, and 4 sets of hairs from the same person as "different," for a false negative rate of 4/9=44 percent. With such high error rates, one would want to study improvements to such systems before putting them into routine practice.

Houck et al. indicate that proficiency testing is conducted regularly for hair experts in crime laboratories.[82] Collaborative Testing Services[83] offers hair and fiber proficiency tests annually. Unfortunately, mass production of test samples such as hair is problematic. Because known samples exhibit a range of characteristics within each of the major and secondary characteristics, it is not possible to provide comparable samples to multiple examiners.

Scientific Interpretation and Reporting of Results

The success of hair analyses to make a positive identification is limited in important ways. Most hair examiners would opine only that hairs exhibiting the same microscopic characteristics "could" have come from a

---

[80] Ibid., p. 179.

[81] Ibid., p. 180.

[82] M.M. Houck, R.E. Bisbing, T.G. Watkins, and R.P. Harman. 2004. Locard exchange: The science of forensic hair comparisons and the admissibility of hair comparison evidence: *Frye* and *Daubert* considered. *Modern Microscopy Journal* Available at www.modernmicroscopy.com/main.asp?article=36&searchkeys=Houck%2BBisbing.

[83] See www.collaborativetesting.com.

particular individual. Moreover, the "best" or most reliable characteristics will vary by case. For example, "color" may be a critical determinant in a case where it is artificial, because that introduces additional independent variables, such as the time since treatment and the actual hair color, while a natural hair might provide less information.

However, several members of the committee have experienced courtroom cases in which, despite the lack of a statistical foundation, microscopic hair examiners have made probabilistic claims based on their experience, as occurred in some DNA exoneration cases in which microscopic hair analysis evidence had been introduced during trial. Aitken and Robertson discuss some probabilistic concepts with respect to hair analysis.[84]

The availability of DNA analysis has lessened the reliance on hair examination. In a very high proportion of cases involving hair evidence, DNA can be extracted, even years after the crime has been committed. Although the DNA extraction may consist of only mitochondrial DNA (mtDNA), such analyses are likely to be much more specific than those conducted on the physical features of hair. For this reason, cases that might have relied heavily on hair examinations have been subjected more recently to additional analyses using DNA.[85] Because of the inherent limitations of hair comparisons and the availability of higher-quality and higher-accuracy analyses based on mtDNA, traditional hair examinations may be presented less often as evidence in the future, although microscopic comparison of physical features will continue to be useful for determining which hairs are sufficiently similar to merit comparisons with DNA analysis and for excluding suspects and assisting in criminal investigations.

Summary Assessment

No scientifically accepted statistics exist about the frequency with which particular characteristics of hair are distributed in the population. There appear to be no uniform standards on the number of features on which hairs must agree before an examiner may declare a "match." In one study of validity and accuracy of the technique, the authors required exact agreement on seven "major" characteristics and at least two agreements among six "secondary" characteristics.[86] The categorization of hair features depends heavily on examiner proficiency and practical experience.

An FBI study found that, of 80 hair comparisons that were "associ-

---

[84] C.G.G. Aitken and J.A. Robertson. 1986. A contribution to the discussion of probabilities and human hair comparisons. *Journal of Forensic Sciences* 32(3):684-689.

[85] M.M. Houck and B. Budowle. 2002. Correlation of microscopic and mitochondrial DNA hair comparisons. *Journal of Forensic Sciences* 47(5):964-967.

[86] R.A. Wickenheiser and D.G. Hepworth. 1990. Further evaluation of probabilities in human hair comparisons. *Journal of Forensic Sciences* 35(6):1323-1329.

ated" through microscopic examinations, 9 of them (12.5 percent) were found in fact to come from different sources when reexamined through mtDNA analysis.[87] This illustrates not only the imprecision of microscopic hair analyses, but also the problem with using imprecise reporting terminology such as "associated with," which is not clearly defined and which can be misunderstood to imply individualization.

In some recent cases, courts have explicitly stated that microscopic hair analysis is a technique generally accepted in the scientific community.[88] But courts also have recognized that testimony linking microscopic hair analysis with particular defendants is highly unreliable.[89] In cases where there seems to be a morphological match (based on microscopic examination), it must be confirmed using mtDNA analysis; microscopic studies alone are of limited probative value. The committee found no scientific support for the use of hair comparisons for individualization in the absence of nuclear DNA. Microscopy and mtDNA analysis can be used in tandem and may add to one another's value for classifying a common source, but no studies have been performed specifically to quantify the reliability of their joint use.

ANALYSIS OF FIBER EVIDENCE

Fibers associated with a crime—including synthetic fibers such as nylon, polyester and acrylic as well as botanical fibers such as ramie or jute, which are common in ropes or twines—can be examined microscopically in the same way as hairs, and with the same limitations. However, fibers also can be analyzed using the tools of analytical chemistry, which provide a more solid scientific footing than that underlying morphological examination. In some cases, clothing and carpets have been subjected to relatively distinctive environmental conditions (e.g., sunlight exposure or laundering agents) that impart characteristics that can distinguish particular items from others from the same manufacturing lot. Fiber examiners agree, however, that none of these characteristics is suitable for individualizing fibers (associating a fiber from a crime scene with one, and only one, source) and that fiber evidence can be used only to associate a given fiber with a class of fibers.[90]

---

[87] Houck and Budowle, op. cit.

[88] E.g., *State v. West*, 877 A.2d 787 (Conn. 2005); *Bookins v. State*, 922 A.2d 389 (Del. Supr, 2007).

[89] See P.C. Giannelli and E. West. 2001. Hair comparison evidence. *Criminal Law Bulletin* 37:514.

[90] See, e.g., R.R. Bresee. 1987. Evaluation of textile fiber evidence: A review. *Journal of Forensic Sciences* 32(2):510-521. See also SWGMAT. 1999. Introduction to forensic fiber examination. *Forensic Science Communications* 1(1). Available at www.fbi.gov/hq/lab/fsc/backissu/april1999/houcktoc.htm, which includes the following summarization in Section 5.4: "It can never be stated with certainty that a fiber originated from a particular textile because

Another type of fiber analysis consists of physically matching two remnants that appear to be torn from one another. By comparing the shapes of the mating edges, and aligning any patterns in the cloth, it can sometimes be possible to associate a fragment with the garment or other item from which it was torn. This is a form of pattern matching, analogous to the matching of shoe and tire prints, but it will not be discussed further here.

Sample Collection and Analysis

The collection of fibers and of a comparison group follows the same procedures as those for mounting hairs. If a macroscopic analysis (e.g., or color, texture, shape) suggests that the two samples appear to be the same, additional procedures such as the following are pursued:

1. Microscopy (reflected light)
2. Polarized light microscopy/fluorescence microscopy
3. Infrared microscopy (to determine man-made fiber composition, such as nylon, polyester)
4. Solubility in a medium
5. Melting point
6. Cross-sectional shape
7. Pyrolysis GC
8. Microspectrophotometry (MSP)
9. Raman spectroscopy

The last of these, Raman spectroscopy, often can provide additional information on polymer chain length (short, medium, long) and branching. Its use in forensic laboratories is rare, although research is under way to develop possible applications. A good overview of fiber evidence is provided by Grieve and Robertson.[91]

Summary Assessment

A group of experienced paint examiners, the Fiber Subgroup of the Scientific Working Group on Materials Analysis (SWGMAT), has produced guidelines,[92] but no set standards, for the number and quality of character-

---

other textiles are produced using the same fiber types and color. The inability to positively associate a fiber to a particular textile to the exclusion of all others, however, does not mean that a fiber association is without value."

[91] M. Grieve and J. Robertson. 1999. *Forensic Examination of Fibres*. London: Taylor and Francis Ltd.

[92] SWGMAT, op. cit. Available at www.fbi.gov/hq/lab/fsc/backissu/april1999/houcktoc. htm.

istics that must correspond in order to conclude that two fibers came from the same manufacturing batch. There have been no studies of fibers (e.g., the variability of their characteristics during and after manufacturing) on which to base such a threshold. Similarly, there have been no studies to inform judgments about whether environmentally related changes discerned in particular fibers are distinctive enough to reliably individualize their source, and there have been no studies that characterize either reliability or error rates in the procedures. Thus, a "match" means only that the fibers could have come from the same type of garment, carpet, or furniture; it can provide only class evidence.

Because the analysis of fibers is made largely through well-characterized methods of chemistry, it would be possible in principle to develop an understanding of the uncertainties associated with those analyses.[93] However, to date, that has not been done. Fiber analyses are reproducible across laboratories because there are standardized procedures for such analyses. Proficiency tests are routinely provided and taken annually, and the reports are available from Collaborative Testing Services.

QUESTIONED DOCUMENT EXAMINATION[94]

Questioned document examination involves the comparison and analysis of documents and printing and writing instruments in order to identify or eliminate persons as the source of the handwriting; to reveal alterations, additions, or deletions; or to identify or eliminate the source of typewriting or other impression marks. Questions about documents arise in business, finance, and civil and criminal trials, and in any matter affected by the integrity of written communications and records. Typical analyses include:

- determining whether the document is the output of mechanical or electronic imaging devices such as printers, copying machines, and facsimile equipment;
- identifying or eliminating particular human or machine sources of handwriting, printing, or typewriting;
- identifying or eliminating ink, paper, and writing instrument;
- establishing the source, history, sequence of preparation, alterations or additions to documents, and relationships of documents;

---

[93] Some relevant questions to be addressed are identified in Bresee, op. cit.

[94] This discussion is primarily based on *Standard Descriptions of Scope of Work Relating to Forensic Document Examiners (American Society for Testing and Materials [ASTM] Designation E 444-98)* (1998), *Standard Guide for Test Methods for Forensic Writing Ink Comparison (ASTM Designation E 1422-01)* (2001), *Standard Guide for Writing Ink Identification (ASTM Designation E 1789-04)* (2004), and *Standard Guide for Examination of Handwritten Items (ASTM Designation E 2290-03)* (2003).

- deciphering and restoring obscured, deleted, or damaged parts of documents;
- recognizing and preserving other physical evidence that may be present in documents; and
- determining the age of a document.

Questioned document examiners are also referred to as forensic document examiners or handwriting experts; questioned document examination includes the field of handwriting identification, while handwriting includes cursive or script style writing, printing by hand, signatures, numerals, or other written marks or signs. Forensic document examination does not involve a study of handwriting that purports to create a personality profile or otherwise analyze or judge the writer's personality or character.

Analyses

Equipment used in questioned document examination includes microscopes and other optical aids, photographic and other imaging devices, and a wide variety of imaging materials adaptable for use with numerous lighting methods, including those involving ultraviolet, visible, and infrared light, and other regions of the electromagnetic spectrum. Software tools recently have become available for the analysis of handwriting.[95] The analysis of papers and inks is similar to other forensic chemistry work. The principal procedures used for ink examination are nondestructive optical examinations and chemical examinations. Optical examinations include those that use visible and alternative light sources—for example, determining whether the class of ink is ballpoint pen; using ultraviolet examination to reveal indications that a document has been stained by chemicals; and employing reflected infrared to observe luminescence at different wavelengths. Chemical examination includes spot testing during which solvents are applied in small amounts to the ink line. For example, ballpoint inks, which are either oil based or glycol based, are highly soluble in pyridine. Inks formulated for fountain pens, porous point pens, and roller pens generally are water soluble in ethanol and water. Indelible markers are solvent based and generally would be soluble in pyridine.

Ink examination can have one of two objectives: class identification— for which the intention is to identify the ink formula or type based on a reference library of samples of inks—and comparison, for which the goal is to compare two ink samples to determine whether they are of common

---

[95] For an overview, see S.N. Srihari and G. Leedham. 2003. A survey of computer methods in forensic document examination. *Proceedings of the 11th International Graphonomics Society Conference*, pp. 278-281. Available at www.ntu.edu.sg/sce/labs/forse/PDF/docExam_7.pdf.

origin. Ink comparisons usually are performed to answer four basic categories of questions: (1) whether an ink is the same (in formula) as that on other parts of the same document or on other documents; (2) whether two writings with similar ink have a common origin (e.g., the same writing instrument or ink well); (3) whether the ink of entries over a period of time is consistent with varying ages or indicates preparation at one time; and (4) whether ink is as old as it purports to be.

Most problems with ink examinations arise from confounding factors that interact with the ink. These can be part of the writing process, such as blotting wet ink; variations in the papers; various forms of contamination on the document; or a combination of these factors. Most ink examinations must be performed on paper and without defacing the handwriting, and this creates a number of sampling and analytical challenges.

The examination of handwritten items typically involves the comparison of a questioned item submitted for examination along with a known item of established origin associated with the matter under investigation. Requirements for comparison are that the writing be of the same type (handwritten/cursive versus hand printed) and that it be comparable text (similar letter/word combinations). Special situations involving unnatural writing are forgery (an attempt to imitate/duplicate the writing of another person) and disguise (an attempt to avoid identification as the writer). The basis for comparison is that handwriting/handprinting/numerals can be examined to obtain writing characteristics (also referred to as features or attributes). The characteristics are further classified into class characteristics (the style that the writer was taught), individual characteristics (the writer's personal style), and gross/subtle characteristics.

Specific attributes used for comparison of handwriting are also referred to as discriminating elements, of which Huber and Headrick have identified 21.[96] Comparisons are based on the high likelihood that no two persons write the same way, while considering the fact that every person's writing has its own variabilities. Thus, an analysis of handwriting must compare interpersonal variability—some characterization of how handwriting features vary across a population of possible writers—with intrapersonal variability—how much an individual's handwriting can vary from sample to sample. Determining that two samples were written by the same person depends on showing that their degree of variability, by some measure, is more consistent with intrapersonal variability than with interpersonal variability. Some cases of forgery are characterized by signatures with too little variability, and are thus inconsistent with the fact that we all have intrapersonal variability in our writing.

---

[96] R.A. Huber and A. M. Headrick. 1999. *Handwriting Identification: Facts and Fundamentals.* Boca Raton, FL: CRC Press.

Scientific Interpretation and Reporting of Results

Terminology has been developed for expressing the subjective conclusions of handwriting comparison and identification, taking into account that there are an infinite number of gradations or opinions toward an identification or elimination. Several scales, such as a five-point scale and a nine-point scale, are used by questioned document examiners worldwide. The nine-point scale is as follows:

1. Identification (a definite conclusion that the questioned writing matches another sample)
2. Strong probability (evidence is persuasive, yet some critical quality is missing)
3. Probable (points strongly towards identification)
4. Indications [that the same person] did [create both samples] (there are a few significant features)
5. No conclusion (used when there are limiting factors such as disguise, or lack of comparable writing)
6. Indications [that the same person] did not [create both samples] (same weight as indications with a weak opinion)
7. Probably did not (evidence is quite strong)
8. Strong probably did not (virtual certainty)
9. Elimination (highest degree of confidence)[97]

Summary Assessment

The scientific basis for handwriting comparisons needs to be strengthened.[98] Recent studies have increased our understanding of the individuality and consistency of handwriting and computer studies[99] and suggest that

---

[97] *Standard Terminology for Expressing Conclusions of Forensic Document Examiners,* *ASTM Designation E 1658-04.*

[98] M. Kam, G. Fielding, and R. Conn. 1997. Writer identification by professional document examiners. *Journal of Forensic Sciences* 42(5):778-786, reports on proficiency tests given to more than 100 questioned document examiners and to a control group of individuals with similar educational backgrounds. Each subject made 144 pair-wise comparisons. Although the study showed that document examiners are much more accurate than lay people in determining whether or not two samples "match" (based on the "identification" and "strong probability" definitions of ASTM standard E1658), professionals nonetheless declared an erroneous match in 6.5 percent of the comparisons. A similar, more recent study, focusing on whether individual signatures were genuine, is reported in J. Sita, B. Found, and D. Rogers. 2002. Forensic handwriting examiners' expertise for signature comparison. *Journal of Forensic Sciences* 47:1117. That study found that professional handwriting examiners erred in 3.4 percent of their judgments.

[99] E.g., S.N. Sargur, S.-H. Cha, H. Arora, and S. Lee. 2002. Individuality of handwriting. *Journal of Forensic Sciences* 47(4):1-17.

there may be a scientific basis for handwriting comparison, at least in the absence of intentional obfuscation or forgery. Although there has been only limited research to quantify the reliability and replicability of the practices used by trained document examiners, the committee agrees that there may be some value in handwriting analysis.

Analysis of inks and paper, being based on well-understood chemistry, presumably rests on a firmer scientific foundation. However, the committee did not receive input on these fairly specialized methods and cannot offer a definitive view regarding the soundness of these methods or of their execution in practice.

ANALYSIS OF PAINT AND COATINGS EVIDENCE

Paint is a suspension of solid pigments in a polymeric binder that, after application by brushing, spraying, dipping, or other means, forms a protective and/or decorative coating. When two objects come in contact with one another and at least one of these objects is painted, a transfer of paint may occur. This transferred paint can be compared to the paint located near the point of damage to determine if the two samples have a common origin. Painted surfaces tend to be repainted over time, providing a characteristic history of layer sequence. Painted surfaces are encountered frequently at crime scenes in the form of vehicles, architectural structures, tools, bicycles, boats, and many other items. The results of the examinations often are valuable both during the investigation and as evidence if a trial results. Paint examinations by their nature can be useful in suggesting possible connections of evidence from the crime scene to its source and therefore are helpful in narrowing or excluding possible witnesses and suspects as well as in providing useful information for investigative leads.

Sample Data and Collection

There are many different types of paint and other coatings, including architectural, vehicular, and marine. Evidence collected from the crime scene may include painted surfaces such as automotive panels, tools, or victims' or suspects' clothing, or spray paint, smears, chips, or flakes. After documentation at the scene, the damaged painted surface is protected and preserved and then submitted to the laboratory. When it is not possible to bring the painted item or a portion of it to the laboratory, paint samples may be removed in such a way that the entire layer sequence is captured intact.

Analyses

The proper recognition and collection of paint evidence at the scene precedes the comparison of evidence occurring at the laboratory. The color, texture, type, layer sequence, and chemical composition of known and questioned paints are compared, and a conclusion is rendered. Additionally, in cases for which no suspect vehicle and questioned paint are available, it may be possible to provide at least an investigative lead based on the color and metallic/nonmetallic type of paint present. If appropriate, the Royal Canadian Mounted Police's PDQ (Paint Data Query) database may be searched, and vehicular information may be provided regarding the possible makes, models, and year range of vehicles that used the questioned paint system.

The examination and comparison of paint evidence requires microscopic and instrumental techniques and methods. The examination of questioned and known samples follows an analytical process that identifies and compares the class (or group) characteristics of the evidence.[100] Occasionally, identifying characteristics exist across edges that allow edge or piece fitting. These characteristics include irregular borders, brush stroke striations, polish mark striations, or surface abrasion markings. When paint fragments physically fit back to a sample from a known source, the fragments are identified as having come from that specific source. Only when physical fitting is possible can an individualized source determination be made

Examiners involved with the analysis of paint evidence in the laboratory typically possess an extensive scientific background, because many of the methods and analyses rely heavily on chemistry.[101] The suggested minimum education requirement is a bachelor's degree in a natural[102,103] or applied science,[104] with many candidates possessing a graduate degree. Coursework needs to include one year (or equivalent) of general chemistry with laboratory, organic chemistry with laboratory, analytical/instrumental analysis, and light microscopy to include basic polarized light microscopy—the latter obtained through structured coursework if it is not available at the graduate or undergraduate level.[105] On-the-job training continues in the laboratory, with its length depending on the examiner's experience. Before examiner trainees can work cases independently, they must observe and

[100] SWGMAT. 1999. Forensic paint analysis and comparison guidelines. *Forensic Science Communications* 1(2). Available at www.fbi.gov/hq/lab/fsc/backissu/july1999/painta.htm.

[101] SWGMAT. 2000. Trace evidence quality assurance guidelines. *Forensic Science Communications* 2(1). Available at www.fbi.gov/hq/lab/fsc/backissu/jan2000/swgmat.htm.

[102] G.S. Anderson (ed.). Canadian Society of Forensic Science. 2007. *CSFS Careers in Forensic Science*, p. 15. Available at www.csfs.ca/contentadmin/UserFiles/File/Booklet2007.pdf.

[103] SWGMAT 2000, op. cit.

[104] Ibid.

[105] Ibid.

work under the supervision of an experienced examiner. The completion of a laboratory's training program in paint analysis can range between 12 to 18 months.[106]

Scientific Interpretation and Reporting of Results

SWGMAT sets guidelines for this field, but it has not recommended report wording, and there are no set criteria for determining a conclusion, although a range of conclusions may be used to show the significance of the examination results. The strength of a conclusion depends on such variables as the number of layers present, the sample condition, and the type of paint (vehicular or structural). Terms such as "matched," "indistinguishable," "consistent," or "similar" are used along with the properties of the paints that were compared in stating the results of the comparison.

If there are no significant differences in the properties compared, the examiners may conclude that the paint or coating samples could have had a common origin. This does not mean they came from the same source to the exclusion of all others, but rather that they may have originated from the same source or from different sources that were painted or coated in the same manner. As the number of different layers associated increases (e.g., multiple different layers on a repainted surface), it may be concluded that it is unlikely that the questioned paint originated from any source other than that of the known paint.

SWGMAT has suggested forensic paint analysis and comparison guidelines[107,108] that discuss the examination procedure and instrumentation options, and ASTM has published the general guidelines.[109] However, neither includes report wording suggestions. Additional work should be done to provide standard language for reporting conclusions and sources of uncertainty. Such work has been completed by working groups for other forensic disciplines. Proficiency testing requirements are agreed upon by the predominant accrediting organization, the American Society of Crime Laboratory Directors-Laboratory Accreditation Board (ASCLD/LAB), which requires testing (internal or external) once per calendar year.

---

[106] Anderson, op. cit.; SWGMAT.

[107] SWGMAT. 1999. Forensic paint analysis and comparison guidelines. *Forensic Science Communications* 1(2). Available at www.fbi.gov/hq/lab/fsc/backissu/july1999/painta.htm.

[108] SWGMAT. 2002. Standard guide for using scanning electron microscopy/X-ray spectrometry in forensic paint examinations. *Forensic Science Communications* 4(4). Available at www.fbi.gov/hq/lab/fsc/backissu/oct2002/bottrell.htm.

[109] Ibid.

Summary Assessment

As is the case with fiber evidence, analysis of paints and coatings is based on a solid foundation of chemistry to enable class identification. Visual and microscopic examinations are typically the first step in a forensic examination of paints and coatings because of the ability to discriminate paints/coatings based on properties determined with these examinations. Several studies have been conducted that included hundreds of random automotive paint samples.[110] These studies have concluded that more than 97 percent of the samples could be differentiated based on microscopic examinations coupled with solubility and microchemical testing. Another study[111] determined that more than 99 percent of 2,000 architectural paint samples could be similarly differentiated. However, the community has not defined precise criteria for determining whether two samples come from a common source class.

ANALYSIS OF EXPLOSIVES EVIDENCE AND FIRE DEBRIS

Explosives evidence encompasses a wide range of materials from unburned, unconsumed powders, liquids, and slurries, to fragments of an explosive device, to objects in the immediate vicinity of an explosion thought to contain residue from the explosive. A typical analytical approach would be to identify the components and construction of an explosive device and conduct an analysis of any unconsumed explosives and residues. In addition to the analysis and identification of low and high explosives, chemical reaction bottle bombs are also analyzed. The scene of an explosion can require special investigative attention. What may appear to be a small piece of scrap metal could in fact be an important piece of the device that caused the explosion. The very nature of an explosion has a direct impact on the quality of evidence recovered. Pristine devices or device fragments, or appreciable amounts of unconsumed explosive material, should not be expected.

Analyses

Generally speaking, laboratories will not accept devices until they have been rendered safe. Examiners involved with the analysis of explosives evidence in the laboratory typically have an extensive scientific background, because the methods used entail a large amount of chemistry and instru-

---

[110] S.G. Ryland and R.J. Kopec. 1979. The evidential value of automobile paint chips. *Journal of Forensic Sciences* 24(1):140-147; J.A. Gothard. 1976. Evaluation of automobile paint flakes as evidence. *Journal of Forensic Sciences* 21(3):636-641.

[111] C.F. Tippet. 1968. The evidential value of the comparison of paint flakes from sources other than vehicles. *Journal of the Forensic Sciences Society* 8(2-3):61-65.

mentation. The Technical Working Group for Fire and Explosives (TW-GFEX), a group of fire debris and explosives examiners, suggests that an explosives examiner be required to possess a bachelor's degree in a natural or applied science, with recommended coursework in chemistry and instrumental analysis.[112] The group also recommends that the examiner complete a training program that includes the analysis of low and high explosives, instruction in the use of instrumentation used in routine analyses, the construction of explosive devices, and participation in a postblast investigation course. Although there is no official certification program for explosives examiners, TWGFEX has devised a suggested training guide. The guide is divided into seven modules, each with a reading list, practical exercises, and methods of evaluation.[113] To ensure that examiners maintain a level of competency, proficiency testing (internal or external) is required by AS-CLD/LAB once per calendar year.[114]

The ultimate goal of an explosives examination is the identification of the explosive material used, whether it is through the analysis of an intact material or of the residue left behind when the material explodes. Intact material lends itself to being more easily identified. The individual components of postblast residue may often be identified (e.g., potassium chloride and potassium sulfate). The training and experience of examiners allows them to deduce what types of explosive material were originally present from possible combinations of explosive materials.

Whether it is a low explosive or high explosive, the analysis of an intact explosive material follows a procedure that begins with a macroscopic and microscopic examination of the material, followed by a burn test, when appropriate. The results of the initial observations will dictate how the rest of the analysis will proceed. Typically it will involve the use of instrumentation that provides both elemental and structural information about the material, such as X-ray diffraction, scanning electron microscope-energy dispersive X-ray analysis, or infrared spectroscopy. TWGFEX has devised guidelines for the analysis of intact explosives that categorize the instruments that can be used based on the level of information they provide.[115] The information gathered, if sufficient, can be useful in identifying the material.

The analysis of postblast explosive residues begins much like the analy-

---

[112] TWGFEX Explosive Examiners Job Description. Undated. Available at http://ncfs.ucf.edu/twgfex/documents.html.

[113] TWGFEX Training Guide for Explosives Analysis Training. Undated. Available at http://ncfs.ucf.edu/twgfex/Documents.html.

[114] American Society of Crime Laboratory Directors International. 2006. *Supplemental Requirements for the Accreditation of Forensic Science Testing Laboratories*, p. 20. See www.ascld-lab.org/international/indexinternational.html.

[115] TWGFEX Recommended Guidelines for Forensic Identification of Intact Explosives. Undated. Available at http://ncfs.ucf.edu/twgfex/documents.html.

sis of intact explosives, with the macroscopic and microscopic analysis of the evidence submitted (whether it is an expended device, fragments of a device, or debris from near the site of the explosion). If no intact explosive material is found, a sequence of extracts may be used to capture any organic and/or inorganic residues present. These extracts are then analyzed employing the same instrumentation used for intact explosives. However, the results produced differ in their specificity, and it is here that the training and expertise of the examiner plays a large role. To interpret the results properly, the examiner must have knowledge of the composition of explosives and the reaction products that form when they explode. Interpretation can be further complicated by the presence of contaminants from, for example, the device or soil.[116]

Examination conclusions for postblast residues range from "the residue present was consistent with an explosive material" to "the residue is only indicative of an explosive" to "no explosive residues were present." TWGFEX recently has developed a set of guidelines for the analysis of postblast explosive residues,[117] but has yet to make any recommendations for report wording.

The examination of fire debris not associated with explosions often aims to determine whether an accelerant was used. To assess the effects of an accelerant, one might design an experiment, under a range of conditions (e.g., wind speed, temperature, presence/absence of other chemicals) with two groups: one in which materials are burned in the presence of an accelerant ("treatment") and one with no accelerant ("control"). The measured outcomes on the burned materials might be measures that characterize the damage patterns (e.g., depth of char, size of bubbles on surfaces). Differences in the ranges of these measurements from the materials in the two groups (treatment versus control) suggest a hypothesis about the effects of an accelerant. Following this exploration, one should design validation studies to confirm that these measures do indeed characterize the differences in materials treated or untreated with an accelerant.

Summary Assessment

The scientific foundations exist to support the analysis of explosions, because such analysis is based primarily on well-established chemistry. As part of the laboratory work, an analyst often will try to reconstruct the bomb, which introduces procedural complications, but not scientific ones.

---

[116] C.R. Midkiff. 2002. Arson and explosive investigation. In: R. Saferstein (ed.). *Forensic Science Handbook*. Vol. 1, 2nd ed. Upper Saddle River, NJ: Prentice Hall.

[117] TWGFEX Recommended Guidelines for Forensic Identification of Post-Blast Explosive Residues. 2007. Available at http://ncfs.ucf.edu/twgfex/action_items.html.

By contrast, much more research is needed on the natural variability of burn patterns and damage characteristics and how they are affected by the presence of various accelerants. Despite the paucity of research, some arson investigators continue to make determinations about whether or not a particular fire was set. However, according to testimony presented to the committee,[118] many of the rules of thumb that are typically assumed to indicate that an accelerant was used (e.g., "alligatoring" of wood, specific char patterns) have been shown not to be true.[119] Experiments should be designed to put arson investigations on a more solid scientific footing.

FORENSIC ODONTOLOGY

Forensic odontology, the application of the science of dentistry to the field of law, includes several distinct areas of focus: the identification of unknown remains, bite mark comparison, the interpretation of oral injury, and dental malpractice. Bite mark comparison is often used in criminal prosecutions and is the most controversial of the four areas just mentioned. Although the identification of human remains by their dental characteristics is well established in the forensic science disciplines, there is continuing dispute over the value and scientific validity of comparing and identifying bite marks.[120]

Many forensic odontologists providing criminal testimony concerning bite marks belong to the American Board of Forensic Odontology (ABFO), which was organized in 1976 and is recognized by the American Academy of Forensic Sciences as a forensic specialty. The ABFO offers board certification to its members.[121]

Sample Data and Collection

Bite marks are seen most often in cases of homicide, sexual assault, and child abuse. The ABFO has approved guidelines for the collection of evidence from bite mark victims and suspected biters.[122] The techniques for obtaining bite mark evidence from human skin—for example, various forms of photography, dental casts, clear overlays, computer enhancement, electron microscopy, and swabbing for serology or DNA—generally are

---

[118] J. Lentini. Scientific Fire Analysis, LLC. Presentation to the committee. April 23, 2007. Available at www7.nationalacademies.org/stl/April%20Forensic%20Lentini.pdf.

[119] NFPA 921 Guide for Explosion and Fire Investigations, 2008 Edition. Quincy, MA: National Fire Protection Association.

[120] E.g., J.A. Kieser. 2005. Weighing bitemark evidence: A postmodern perspective. *Journal of Forensic Science, Medicine, and Pathology* 1(2):75-80.

[121] American Board of Forensic Odontology at www.abfo.org.

[122] Ibid.

well established and relatively noncontroversial. Unfortunately, bite marks on the skin will change over time and can be distorted by the elasticity of the skin, the unevenness of the surface bite, and swelling and healing. These features may severely limit the validity of forensic odontology. Also, some practical difficulties, such as distortions in photographs and changes over time in the dentition of suspects, may limit the accuracy of the results.[123]

Analyses

The guidelines of the ABFO for the analysis of bite marks list a large number of methods for analysis, including transillumination of tissue, computer enhancement and/or digitalization of the bite mark or teeth, stereomicroscopy, scanning electron microscopy, video superimposition, and histology.[124] The guidelines, however, do not indicate the criteria necessary for using each method to determine whether the bite mark can be related to a person's dentition and with what degree of probability. There is no science on the reproducibility of the different methods of analysis that lead to conclusions about the probability of a match. This includes reproducibility between experts and with the same expert over time. Even when using the guidelines, different experts provide widely differing results and a high percentage of false positive matches of bite marks using controlled comparison studies.[125]

No thorough study has been conducted of large populations to establish the uniqueness of bite marks; theoretical studies promoting the uniqueness theory include more teeth than are seen in most bite marks submitted for comparison. There is no central repository of bite marks and patterns. Most comparisons are made between the bite mark and dental casts of an individual or.individuals of interest. Rarely are comparisons made between the bite mark and a number of models from other individuals in addition to those of the individual in question. If a bite mark is compared to a dental cast using the guidelines of the ABFO, and the suspect providing the dental cast cannot be eliminated as a person who could have made the bite, there is no established science indicating what percentage of the population or subgroup of the population could also have produced the bite. This follows from the basic problems inherent in bite mark analysis and interpretation.

As with other "experience-based" forensic methods, forensic odontology suffers from the potential for large bias among bite mark experts in evaluating a specific bite mark in cases in which police agencies provide the suspects for comparison and a limited number of models from which

---

[123] Rothwell, op. cit.

[124] American Board of Forensic Odontology, op. cit.

[125] Bowers, op. cit.

to choose from in comparing the evidence. Bite marks often are associated with highly sensationalized and prejudicial cases, and there can be a great deal of pressure on the examining expert to match a bite mark to a suspect. Blind comparisons and the use of a second expert are not widely used.

Scientific Interpretation and Reporting of Results

The ABFO has issued guidelines for reporting bite mark comparisons, including the use of terminology for conclusion levels, but there is no incentive or requirement that these guidelines be used in the criminal justice system. Testimony of experts generally is based on their experience and their particular method of analysis of the bite mark. Some convictions based mainly on testimony by experts indicating the identification of an individual based on a bite mark have been overturned as a result of the provision of compelling evidence to the contrary (usually DNA evidence).[126]

More research is needed to confirm the fundamental basis for the science of bite mark comparison. Although forensic odontologists understand the anatomy of teeth and the mechanics of biting and can retrieve sufficient information from bite marks on skin to assist in criminal investigations and provide testimony at criminal trials, the scientific basis is insufficient to conclude that bite mark comparisons can result in a conclusive match. In fact, one of the standards of the ABFO for bite mark terminology is that, "Terms assuring unconditional identification of a perpetrator, or without doubt, are not sanctioned as a final conclusion."[127]

Some of the basic problems inherent in bite mark analysis and interpretation are as follows:

(1) The uniqueness of the human dentition has not been scientifically established.[128]

(2) The ability of the dentition, if unique, to transfer a unique pattern to human skin and the ability of the skin to maintain that uniqueness has not been scientifically established.[129]

    i. The ability to analyze and interpret the scope or extent of distortion of bite mark patterns on human skin has not been demonstrated.

    ii. The effect of distortion on different comparison techniques is not fully understood and therefore has not been quantified.

---

[126] Bowers, op. cit.

[127] American Board of Forensic Odontology, op. cit.

[128] Senn, op. cit.

[129] Ibid.

(3) A standard for the type, quality, and number of individual charac-
teristics required to indicate that a bite mark has reached a thresh-
old of evidentiary value has not been established.

Summary Assessment

Despite the inherent weaknesses involved in bite mark comparison, it
is reasonable to assume that the process can sometimes reliably exclude
suspects. Although the methods of collection of bite mark evidence are
relatively noncontroversial, there is considerable dispute about the value
and reliability of the collected data for interpretation. Some of the key ar-
eas of dispute include the accuracy of human skin as a reliable registration
material for bite marks, the uniqueness of human dentition, the techniques
used for analysis, and the role of examiner bias.[130] The ABFO has devel-
oped guidelines for the analysis of bite marks in an effort to standardize
analysis,[131] but there is still no general agreement among practicing forensic
odontologists about national or international standards for comparison.

Although the majority of forensic odontologists are satisfied that bite
marks can demonstrate sufficient detail for positive identification,[132] no
scientific studies support this assessment, and no large population studies
have been conducted. In numerous instances, experts diverge widely in their
evaluations of the same bite mark evidence,[133] which has led to questioning
of the value and scientific objectivity of such evidence.

Bite mark testimony has been criticized basically on the same grounds
as testimony by questioned document examiners and microscopic hair ex-
aminers. The committee received no evidence of an existing scientific basis
for identifying an individual to the exclusion of all others. That same find-
ing was reported in a 2001 review, which "revealed a lack of valid evidence
to support many of the assumptions made by forensic dentists during bite
mark comparisons."[134] Some research is warranted in order to identify
the circumstances within which the methods of forensic odontology can
provide probative value.

---

[130] Ibid.

[131] American Board of Forensic Odontology, op. cit.

[132] I.A. Pretty. 2003. A Web-based survey of odontologists' opinions concerning bite mark
analyses. *Journal of Forensic Sciences* 48(5):1-4.

[133] C.M. Bowers. 2006. Problem-based analysis of bite mark misidentifications: The role of
DNA. *Forensic Science International* 159 Supplement 1:s104-s109.

[134] I.A. Pretty and D. Sweet. 2001. The scientific basis for human bitemark analyses—A
critical review. *Science and Justice* 41(2):85-92. Quotation taken from the abstract.

BLOODSTAIN PATTERN ANALYSIS

Understanding how a particular bloodstain pattern occurred can be critical physical evidence, because it may help investigators understand the events of the crime. Bloodstain patterns occur in a multitude of crime types—homicide, sexual battery, burglary, hit-and-run accidents—and are commonly present. Bloodstain pattern analysis is employed in crime reconstruction or event reconstruction when a part of the crime scene requires interpretation of these patterns.

However, many sources of variability arise with the production of bloodstain patterns, and their interpretation is not nearly as straightforward as the process implies. Interpreting and integrating bloodstain patterns into a reconstruction requires, at a minimum:

- an appropriate scientific education;
- knowledge of the terminology employed (e.g., angle of impact, arterial spurting, back spatter, castoff pattern);
- an understanding of the limitations of the measurement tools used to make bloodstain pattern measurements (e.g., calculators, software, lasers, protractors);
- an understanding of applied mathematics and the use of significant figures;
- an understanding of the physics of fluid transfer;
- an understanding of pathology of wounds; and
- an understanding of the general patterns blood makes after leaving the human body.

Sample Data and Collection

Dried blood may be found at crime scenes, deposited either through pooling or via airborne transfer (spatter). The patterns left by blood can suggest the kind of injury that was sustained, the final movements of a victim, the angle of a shooting, and more. Bloodstains on artifacts such as clothing and weapons may be crucial to understanding how the blood was deposited, which can indicate the source of the blood. For example, a stain on a garment, such as a shirt, might indicate contact between the person who wore the shirt and a bloody object, while tiny droplets of blood might suggest proximity to a violent event, such as a beating.

Analyses

Bloodstain patterns found at scenes can be complex, because although overlapping patterns may appear simple, in many cases their interpreta-

tions are difficult or impossible. [135,136] Workshops teach the fundamentals of basic pattern formation and are not a substitute for experience and experimentation when applying knowledge to crime reconstruction.[137] Such workshops are more aptly applicable for the investigator who needs to recognize the importance of these patterns so that he or she may enlist the services of a qualified expert. These courses also are helpful for attorneys who encounter these patterns in the course of preparing a case or when preparing to present testimony in court.

Although there is a professional society of bloodstain pattern analysts, the two organizations that have or recommend qualifications are the IAI and the Scientific Working Group on Bloodstain Pattern Analysis (SWGSTAIN). SWGSTAIN's suggested requirements for practicing bloodstain pattern analysis are outwardly impressive, as are IAI's 240 hours of course instruction. But the IAI has no educational requirements for certification in bloodstain pattern analysis.[138] This emphasis on experience over scientific foundations seems misguided, given the importance of rigorous and objective hypothesis testing and the complex nature of fluid dynamics. In general, the opinions of bloodstain pattern analysts are more subjective than scientific. In addition, many bloodstain pattern analysis cases are prosecution driven or defense driven, with targeted requests that can lead to context bias.

Summary Assessment

Scientific studies support some aspects of bloodstain pattern analysis. One can tell, for example, if the blood spattered quickly or slowly, but some experts extrapolate far beyond what can be supported. Although the trajectories of bullets are linear, the damage that they cause in soft tissue and the complex patterns that fluids make when exiting wounds are highly variable. For such situations, many experiments must be conducted to determine what characteristics of a bloodstain pattern are caused by particular actions during a crime and to inform the interpretation of those causal links and

---

[135] H.L. MacDonell. 1997. Bloodstain Patterns. Corning, NY: Laboratory of Forensic Science; S. James. 1998. *Scientific and Legal Applications of Bloodstain Pattern Interpretation.* Boca Raton, FL: CRC Press; P. Pizzola, S. Roth, and P. DeForest. 1986. Blood drop dynamics–II. *Journal of Forensic Sciences* 31(1): 36-49.

[136] Ibid.; R.M. Gardner. 2004. *Practical Crime Scene Processing and Investigation.* Boca Raton, FL: CRC Press; H.C. Lee; T. Palmbach and M.T. Miller. 2005. *Henry Lee's Crime Scene Handbook.* Burlington, MA: Elsevier Academic Press, pp. 281-298.

[137] W.J. Chisum and B.E. Turvey. 2007. *Crime Reconstruction.* Burlington, MA: Elsevier Academic Press.

[138] See "Bloodstain Pattern Examiner Certification Requirements." Available at theiai.org/certifications/bloodstain/requirements.php.

their variabilities. For these same reasons, extra care must be given to the way in which the analyses are presented in court. The uncertainties associated with bloodstain pattern analysis are enormous.

AN EMERGING FORENSIC SCIENCE DISCIPLINE:
DIGITAL AND MULTIMEDIA ANALYSIS

The analysis of digital evidence deals with gathering, processing, and interpreting digital evidence, such as electronic documents, lists of phone numbers and call logs, records of a device's location at a given time, e-mails, photographs, and more. In addition to traditional desktop and laptop computers, digital devices that store data of possible value in criminal investigations include cell phones, GPS devices, digital cameras, personal digital assistants (PDAs), large servers and storage devices (e.g., RAIDS and SANS), video game consoles (e.g., PlayStation and Xbox), and portable media players (e.g., iPods). The storage media associated with these devices currently fall into three broad categories. The first, magnetic memory, includes hard drives, floppy discs, and tapes. The second, optical memory, includes compact discs (CDs), and digital versatile discs (DVDs). The third, electrical storage, includes USB flash drives, some memory cards, and some microchips. These items are the most commonly encountered in criminal and counterintelligence matters, but laboratories have been asked to examine such items as scuba dive watches in death investigations and black boxes in aircraft mishaps.

The proliferation of computers and related devices over the past 30 years has led to significant changes in and the expansion of the types of criminal activities that generate digital evidence. Initially, computers were either the weapon or the object of the crime. In the early days, most computer crime involved manipulating computer programs of large businesses in order to steal money or other resources. As computers became more popular, they became storage containers for evidence. Drug dealers, book makers, and white collar criminals began to keep computerized spreadsheets detailing their transactions. Digital cameras and the Internet have made child pornography increasingly available, and computers act as a digital file cabinet to hold this contraband material. Finally, digital media have become witnesses to daily activities. Many individuals have two cell phones with text messaging and/or e-mail capability, several computers, a home alarm system, a GPS in the car, and more; even children often possess some subset of these items. Workplaces use magnetic card readers to permit access to buildings. Most communication involves some kind of computer, and by the end of each day, hundreds of megabytes of data may have been generated about where individuals have been, how fast they got there, to whom they spoke, and even what was said. Suicide notes are written on

computers. Sexual predators stalk their victims online via e-mail, chat, and instant messaging. Even get-away cars are equipped with GPS devices. Finally, computer systems have become (with ever-increasing frequency) the victims of unauthorized control or intrusions. These intrusions often result in the manipulation of files and the exfiltration of sensitive information. In addition, computers in automobiles that track speed, breaking, and turning are valuable in accident reconstruction. As a result, almost every crime could have digital evidence associated with it.

Sample Data and Collection

The best practices for the collection of digital evidence most often call for the person at the scene to disconnect the power cord for the computer and related peripheral equipments (e.g., monitor, printer) and seize these items, as well as any loose storage media such as thumb drives and CDs. This method works well in most cases. However, some data (like recently typed passwords, malicious programs, and active communication programs) are volatile and are stored in the electronic chips of the system. In these circumstances, this information is lost when the device is turned off. In intrusion investigations or in cases in which encryption software is being used, this volatile information could be the key to a successful analysis and prosecution.[139]

Recognizing potential sources of digital evidence is also an ongoing challenge. Investigators are likely to seize a desktop computer but walk past a PlayStation. Thumb drives can be fashioned to look like a pocket knife, writing pen, or even a piece of sushi. Cell phones and wireless Internet capability present another challenge: If these devices are turned on while in law enforcement custody, they could be remotely accessed and altered by a suspect.

Analyses

The typical approach to examining a computer involves two main phases. The first is the imaging phase. During this process, the storage device (most often a hard drive) is fitted with an appliance that prevents any new information from being written. Then, all of the data are copied to a new blank hard drive. The copy is compared with the original, most often by using a mathematical algorithm called Message Digest–5, otherwise known as MD5 Hash. The MD5 Hash value gives a unique series of numbers and letters for every file. In the examination phase, this forensi-

---

[139] See W.G. Kruse and J.G. Heiser. 2001. *Computer Forensics: Incident Response Essentials.* Boston: Addison-Wesley.

cally sound copy is examined for saved computer files with probative value. These so-called logical files often are pictures, documents, spreadsheets, and e-mail files that have been saved by the user in various folders or directories. Logical files are patent evidence. Next, the forensic copy is examined for files that have previously been deleted. The computer files are sometimes called physical, because the data are physically present on the hard drive but they are not logically available to the computer operating system. Such files constitute latent evidence.

Finally, system files that are created and saved by the operating system are examined. These files are analogous to a surveillance tape that shows programs that were running on the computer and files that were changed. The goal of most of these examinations is to find files with probative information and to discover information about when and how these files came to be on the computer.[140]

Digital evidence has undergone a rapid maturation process. This discipline did not start in forensic laboratories. Instead, computers taken as evidence were studied by police officers and detectives who had some interest or expertise in computers. Over the past 10 years, this process has become more routine and subject to the rigors and expectations of other fields of forensic science. Three holdover challenges remain: (1) the digital evidence community does not have an agreed certification program or list of qualifications for digital forensic examiners; (2) some agencies still treat the examination of digital evidence as an investigative rather than a forensic activity; and (3) there is wide variability in and uncertainty about the education, experience, and training of those practicing this discipline.

A publication of the Department of Justice Computer Crime and Intellectual Property Section, *Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations,*[141] describes the challenging legal issues surrounding the examination of digital evidence. For example, sometimes the courts have viewed computers as a piece of evidence that is sent to a laboratory for forensic examination, and as having no special legal constraints, while other times, the courts have viewed computers as a virtual room or filing cabinet.[142] For the latter cases, a warrant must be

---

[140] See E. Casey. 2004. *Digital Evidence and Computer Crime.* San Diego, CA: Academic Press; E. Casey. 2001. *Handbook of Computer Crime Investigation: Forensic Tools & Technology.* San Diego, CA: Academic Press; B. Carrier. 2005. *File System Forensic Analysis.* Boston: Addison-Wesley; S. Anson and S. Bunting. 2007. *Mastering Windows Network Forensics and Investigation.* Indianapolis: Sybex; and H. Carvey and D. Kleiman. 2007. *Windows Forensic Analysis.* Burlington: Syngress.

[141] Available at www.usdoj.gov/criminal/cybercrime/s&smanual2002.htm.

[142] See, e.g., G.R. McLain, Jr., 2007. *United States v. Hill:* A new rule, but no clarity for the rules governing computer searches and seizures. *George Mason Law Review* 14(4):1071-1104; D. Regensburger, B. Bytes, and B. Bonds. 2007. An exploration of the law concerning

obtained that specifies how the examination will be conducted and which files can be recovered before the electronic device can be examined.

Finally, the analysis of digital evidence differs from other forensic science disciplines because the examination generates not only a forensic report, but also brings to light documents, spreadsheets, and pictures that may have probative value. Different agencies have handled these generated files in different ways: Some treat them as exhibits, while others treat them as derivative evidence that requires a chain of custody and special protection.

A growing number of colleges and universities offer courses in computer security and computer forensics. Still, most law enforcement agencies are understaffed in trained computer security experts.

CONCLUSIONS

The term "forensic science" encompasses a broad range of disciplines, each with its own set of technologies and practices. Wide variability exists across forensic science disciplines with regard to techniques, methodologies, reliability, error rates, reporting, underlying research, general acceptability, and the educational background of its practitioners. Some of the forensic science disciplines are laboratory based (e.g., nuclear and mitochondrial DNA analysis, toxicology, and drug analysis); others are based on expert interpretation of observed patterns (e.g., fingerprints, writing samples, toolmarks, bite marks, and specimens such as fibers, hair, and fire debris). Some methods result in class evidence and some in the identification of a specific individual—with the associated uncertainties. The level of scientific development and evaluation varies substantially among the forensic science disciplines.

---

the search and seizure of computer files and an analysis of the Ninth Circuit's decision in *United States v. Comprehensive Drug Testing, Inc. Journal of Criminal Law and Criminology* 97(4)1151-1208.